IMPACT

of science on society

THE UNIVERSITY OF MICHIGAN
MAY 6 1960
PERIODICAL READING ROOM

Vol. X (1960), No. 1

SPECIAL NUMBER ON INFORMATION PROCESSING

Computers and Automation by E. L. HARDER
Machine Translation and the Human Being by D. J. PANOV
Machine Translation of Languages: Research and Organizational Problems by E. DELAVENAY
Current Problems and Recent News

IMPACT *of Science on Society*
quarterly

Annual subscription:
[A] $1.75; 9/6 (stg.); 4,50 NF
Per copy:
[A] $0.50; 2/6 (stg.); 1,25 NF

Any of the National Distributors listed at the end of this number will accept subscriptions; rates in currencies other than the above will be supplied on application to the National Distributor in the country concerned.
When notifying change of address please enclose last wrapper or envelope.

The articles appearing in *Impact* express the views of their authors, and not necessarily those of Unesco.

© Unesco 1960
Printed in France
NS.59/I.34/A

UNITED NATIONS EDUCATIONAL, SCIENTIFIC
AND CULTURAL ORGANIZATION
Place de Fontenoy, Paris-7e (France)

Contents

Impact, Vol. X (1960), No. 1

CB
478
.A1
I34

COMPUTERS AND AUTOMATION
by E. L. HARDER 3

In this article, the author describes the increasingly important part played by electronic computers in industry and the business world. He makes a critical study of the automation of manufacturing and management processes, thus giving the reader a preview of the scientific organization of the world of tomorrow.

MACHINE TRANSLATION AND THE HUMAN BEING
by D. J. PANOV. 16

The problem of translation is one of the most serious with which present-day mankind has to deal, since upon its successful solution depends our awareness of what is going on in the world, our understanding of other nations, the dissemination of knowledge and the possibility of establishing international co-operation in every sphere of intellectual activity. A very attractive way of dealing with the problem is presented by the use, for translation purposes, of electronic machines, similar to those employed for mathematical computation.

MACHINE TRANSLATION OF LANGUAGES: RESEARCH AND ORGANIZATIONAL PROBLEMS
by E. DELAVENAY 26

Within a circumscribed field, machine translation of languages by means of electronic computers of the 'ordinating' type is already possible, as the history of research in this connexion demonstrates. To raise machine translation to the level of operational practice, enlarge its range and take advantage of the theoretical data it can provide, it will be necessary to organize a double cycle of work: scientific training in the discipline of applied linguistics for the men who will work with the machines; and formulation of programmes for the use of the machines. These are eminently team-work tasks of international scope, in a branch of humanistic studies in which the observation methods of the natural sciences have an essential part to play.

CURRENT PROBLEMS AND RECENT NEWS

The Scientific Attitude: a possible misunderstanding
by P. AUGER 45

The essential phenomena, in modern science, are those which take place at microscopic scale, between nuclei, atoms and molecules. Thus it is in the field of physics and chemistry that the solutions of a growing number of the problems of nature are being discovered. Direct actions at human scale, which played such an important role in the theoretical views of past ages, have to be resolved now in a sum of elementary actions, and the causal chain of successive events which joins without any gap in space or time the initial fact to the final result, must be at least conceivable. Any type of action which does not conform to this requirement should be rejected by the scientific mind. These considerations should help to clarify the misunderstandings which take place so frequently between the scientists and those persons who, in good faith, would like to introduce into science observations or theories which have no place there.

International Conference on Information Processing
by S. DE PICCIOTTO 53

Electronic computers are occupying an increasingly important place in science, technology and industry. It was this consideration which led Unesco to organize a major international conference on the subject at its Headquarters in Paris from 15 to 20 June 1959. An account of the conference proceedings is presented here for the information of the layman.

CONTENTS OF PRECEDING ISSUES

Vol. IX, No. 3, 1959

Grafting for Productivity in Plants and Animals, by PIERRE BOITEAU.
The Development of Optics and its Impact on Society, by VASCO RONCHI.
Teaching the History of Science, by RENÉ TATON.
Current Problems and the 1959-65 Seven-Year Plan in the USSR, by NORAIR SISSAKIAN.
The International Conference on Scientific Information, by PAUL BOQUET.

Vol. IX, No. 4, 1959

How Life on Earth began, by A. I. OPARIN.
The Place of Laboratory Animals in the Scientific Life of a Country, by W. LANE-PETTER.
The Place of Operational Research in the Development of Modern Society, by J. LESOURNE.
Science in Modern India, by Dr. A. RANGANANTHAN.
Problems of Waste Disposal in the Wide-scale Use of Radio-isotopes, by W. G. MARLEY.

COMPUTERS AND AUTOMATION

by
E. L. HARDER

E. L. Harder is the Director of the Analytical Department of the Westinghouse Electric Corporation at East Pittsburgh, Pennsylvania, USA. This article is based on a lecture presented at the International Conference on Information Processing, held on 18 June 1959 in Paris.

There is no doubt that the rapid strides made in computer technology are giving a new dimension to the term 'automation'. It is useful to outline this phase of one of the great technologies of our time—the relationship between computers and automation—for the purpose of computers and information-processing machines is to reduce man's work and to amplify through automation the work he performs in producing the necessities and luxuries of life.

All man's work can be divided into two general categories, it is either physical or it is mental or clerical work; and the development of computers and information-processing machines can be traced through their association with both types of work.

In the category of mental work, early aids to calculation were the desk calculator, the slide rule, and the cash register. Next came the business machine, popularly known as tabulating equipment, capable of performing, very rapidly, simple calculations on large numbers of similar documents of the kind frequently used in business and government operations. Next came the large-scale, internally programmed, electronic computer which, for the first time, offered tremendously lowered unit cost for calculation. A calculation which only a few years before required two weeks using desk calculators and which cost $300 could now be made for 3 cents on a modern high-speed computer, a reduction of 10,000 to 1. The explosive effect of this tremendously lowered unit cost of calculation can well be imagined. In business it has been put to work in the programming of complete sequences of business calculations. In engineering it has been used for the solution of many problems that it was previously impossible, or hopelessly uneconomic, to solve by any known means of calculation. But, also, it has been used for reducing the routine work of design to a mechanical operation, releasing valuable technical manpower for developing and improving the methods of design.

The internally programmed computer brought with it, however, a great problem, that of communicating with the machine or 'programming'. This

work accounted for about half the total cost of operating a computing installation, since the early computers had to be directed in minute detail as to each infinitesimal step to be taken in the solution of a problem. It is in this area that great strides are now being made through the development of programming aids and automatic programming, to which we shall refer later.

Turning now to man's physical work, we see how computing aids have come to be used in connexion with his tools. Far back in history, man started using the tool, and he himself supplied both the intelligence and the power for its operation. Next he harnessed the forces of nature, first animals, then hydro power and the fossil fuels, and now the atom, and through electricity brought power to the tool; but here also man still provided practically all the intelligence and the direction. Next came the automatically controlled power tool with much of the intelligence for the immediate operation built into the tool, usually in the form of analogue computing devices, servo-mechanisms, and feedback control. However, most of the planning, sequencing, optimizing and analysing of the results leading to corrective action was left with man.

Now we are coming to the era of completely sequenced automatic controls in which more sophisticated computing functions are being performed. The 'intelligence' or computer part of the machine is beginning to emerge as a distinct organ separate from the power-handling facilities. Here we may cite the completely sequenced programme control of a steel rolling mill, the economic dispatching computer control for an electric power system or the optimizing controls for the process industry.

It requires little imagination to visualize the enormous possibilities that remain practically unexplored in this direction, as the power of mechanized logic and computing is brought to bear on the sequencing and optimizing of controlled processes of all types.

Let us follow these two main channels of computer application and dwell briefly on a few of the significant developments which, taken in total, help to convey a picture of the extent and significance of this revolution that is taking place in the processing of information. This revolution in the automating of mental and clerical work is no less significant than was the industrial revolution in the automating of much of man's physical work.

PROGRESS IN COMPUTERS

Turning first to the developments in the computer itself, the curve in Fig. 1 shows how the successive computer developments over the last several years have increasingly lowered the unit cost of calculation by a factor of some 10,000 to 1 up to the present, and further reductions which will take place in the future. In the light of the developments at present in prospect there

is no reason whatever to expect that this curve will flatten off before further decades have passed during which greater computer power per unit cost will have been made available, with further radical extensions in the areas to which computers may be economically employed.

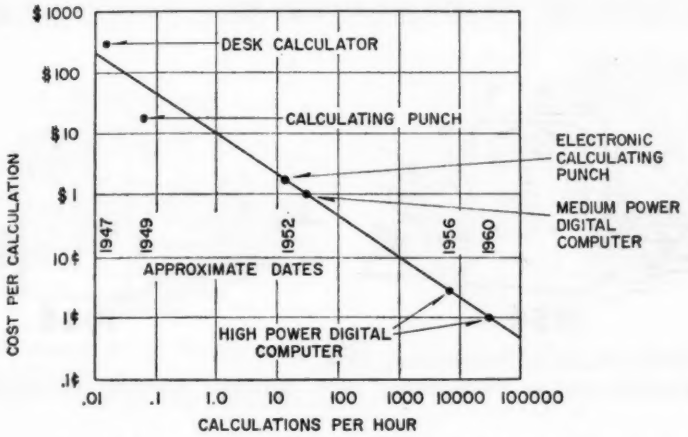


Fig. 1. Computation Cost versus Computer Speed.

PROGRESS IN AUTOMATIC PROGRAMMING

There has always been much speculation regarding the relationship of the computer to the human brain, and opinions differ as to whether the computer should be likened to a giant brain or a giant moron; the latter opinion was held by the early programmers who realized what detailed instructions had to be given even though the computer operated extremely fast. However, the progress in automatic programming is enabling us to communicate more readily with the computer, and a computer coupled with its system of programmes constitutes an extremely intelligent device which, in the areas in which it has been taught or programmed, can certainly outperform the best brains by a large factor.

In order to demonstrate the progress in automatic programming we must first describe the situation at the time of the earlier computers. With the large-scale computer of 1954 (Fig. 2) where the computer had to be directed in every step of the calculation in great detail, the programmer had a tremendous job to do. The answers were not turned out very fast and, as the illustration shows, the computer is somewhat unhappy about this.

By 1956, great progress had been made in automatic programming; to give an example, for the IBM-704 computer this took the form of a computer programme called SAP, meaning Symbolic Assembly Programme. SAP may be viewed as a gremlin located in the computer (Fig. 3). Actually, it is a programme on magnetic tapes in the computer. When the programmer writes a new programme for a specific job, he now brings it to a computer

which already contains the SAP programme and SAP performs the following operations which formerly were the work of the programmer. It concerts the simple symbolic names for the commands into the decimal or binary machine codes as required. It brings into the programme any needed routines from the library tape, allocates the memory space, inserts correct addresses

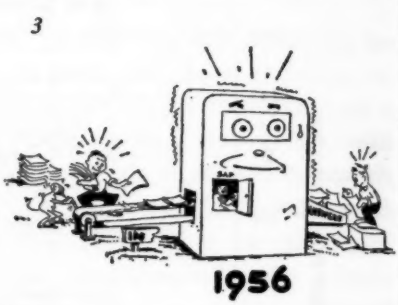


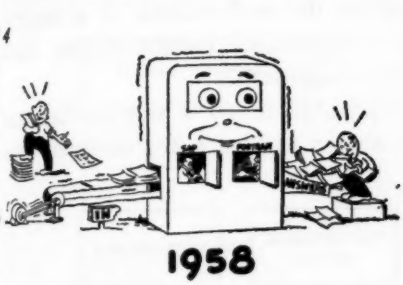
Fig. 2. Developments in Programming 1954.
Fig. 3. Developments in Programming 1956—Symbolic Assembly Programme.

in all the instructions and detects common programmer's errors. It converts from our decimal number system to the binary number system if that is used. Finally, it condenses the programme to a tape or compact card form. With this tremendous aid to programming, known as automatic programming, a great deal less effort was required for writing programmes for particular problems. Larger problems could be approached; the answers came out very much more rapidly; the programmer's work was lightened; and the computer is somewhat more happy about this.

During the next two years, further advances in automatic programming added a second gremlin to the computer, known as the compiler programme. Fortran for the IBM-704 (Fig. 4) is an example of such a programme and is a great advance over the assembly programme called SAP. In using an assembly programme, the programmer must write each step, except for subroutines, although in abbreviated form. With a compiler, this is no longer necessary. He simply writes the mathematical statement for the programme to be prepared and the compiler, i.e. the Fortran programme, does the rest. It translates the mathematical statements into a machine programme with an efficiency of 5 or 10 to 1 as compared with writing out each step, as was necessary two years earlier. Fortran is designed primarily for scientific or engineering problems although similar computing aids are being developed for business and data processing problems. With Fortran the programmer's job is very much lightened, the answers flow out at an increased rate, and, as you see, the computer is quite happy about this.

Fig. 5 illustrates the hope of the programmers, that an additional gremlin can be put in the computer to take over the rest of their work. No doubt this millenium is still some distance off; although progress continues in this

area. Specialized languages are being written for particular problems, for example, the giving of directions to a machine tool. Also, a language was devised recently for point mechanics, enabling any problem in masses, springs, levers, torques, and friction to be simply described to the computer. Work is being done on a universal language which will provide compati-



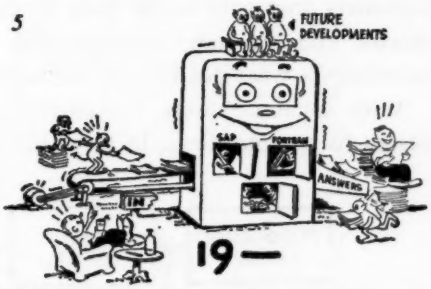


Fig. 4. Developments in Programming 1958—Compiler Programme, Fortran, Formula Translation.
Fig. 5. Future Programming Developments 19—.

bility between computers, making it possible to run problems on one computer that have been programmed for another. In business, the concept of report generator implies the ability to specify conveniently and quickly a new derivation from the data needed for certain management decisions, an advantage which would be quite impossible without a highspeed computer.

ENGINEERING AND SCIENTIFIC CALCULATIONS

Turning to engineering calculations, we find that computers are used most effectively in the design of apparatus and systems. One of the simplest design calculations is the performance calculation (Fig. 6) in which the designer

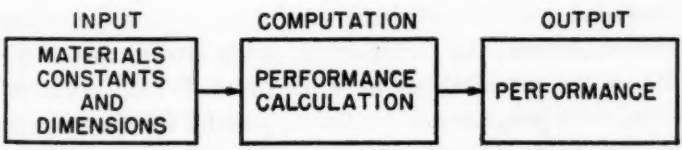


Fig. 6. PERFORMANCE CALCULATION

specifies the shape and materials of the machine or system and programmes the formulae for calculating the resulting performance. These formulae are usually available so that it is simply a question of substituting the computer for a calculation sheet. Simple as it seems, this possibility of calculating the performance through the use of a computer is a tremendous aid to the

designer since the computer can make the calculations hundreds of times faster. The designer can therefore calculate problems of greater complexity than would otherwise be possible and can try out many alternatives, thus improving the design. For the more complicated engineering systems and devices, this is as far as any one has gone and, in fact, we consider ourselves extremely fortunate to be able to calculate the performance of a nuclear reactor and the natural frequencies of a system of turbine blades. Such calculations were not feasible without the computer.

With some products, however, it is possible to go a step further and design to specifications. In this case, as shown in Fig. 7, the designer feeds

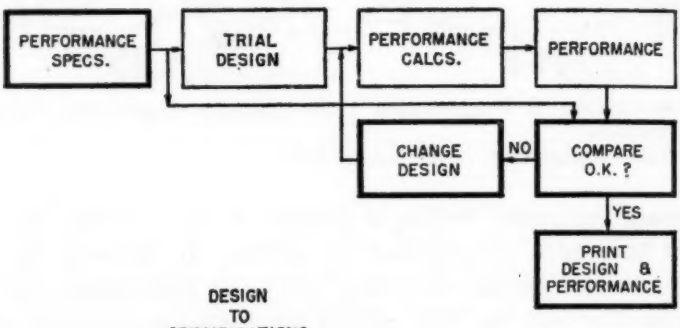


Fig. 7. DESIGN TO SPECIFICATIONS

to the computer the specifications for the particular product to be designed. Usually this is one of a line of products, such as a motor or a transformer, for which the individual unit has not been designed but for which the general form and method of design has been worked out in advance. He also feeds in enough empirical information or rough formulae so that the computer can determine an initial trial design. The computer then makes the performance calculations as before, but when it gets the performance for the trial design it compares these with the specified performance. If these do not correspond, and are either too low or too high, the computer makes changes and recalculates the performance, going around this loop of calculations until all the specifications have been met. It then prints out the resulting design. Such programmes are few at present, for they can be handled only by skilled designers who know what to change in order to be sure of converging on a design which meets all the specifications. Only a skilled designer knows in what order to make these changes so that having met nine of the specifications, he does not come to a tenth, which would require all the work to be thrown out and a new start made. Transformers, certain motors, and a number of other products are now designed to specifications by computers.

The next step along this line (Fig. 8) is the optimizing design programme, a truly monumental accomplishment. Here, the computer does all that it

did before but it is not content to stop when it reaches a design which meets all the specifications. It ranges further over permissible variations in the form or material of the device, generating not one but many designs, all of which meet all the specifications, and amongst which it seeks the optimum by some criterion such as cost or weight, or whatever may be most impor-

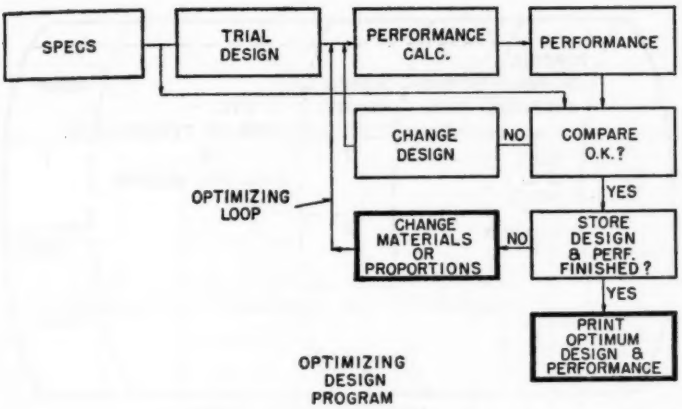


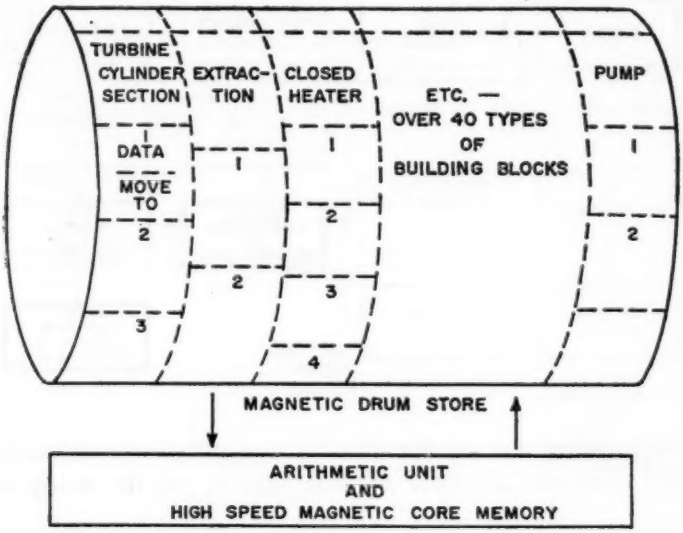
Fig. 8.

tant. A single programme of this type for designing large induction motors required 10 man-years of effort and has been in use for nearly two years now.

In order to simplify the programming of a problem relating to a large system, the building block approach is used. For example, a large steam system consists of many boilers, turbine cylinders, heaters, pumps and other devices. In all, there are more than forty different kinds of devices involved and several of each may be involved in a single system. The principal problem which must be solved in properly applying all this equipment is called the heat balance problem. It consists of striking a balance between the heat of the steam at various points and the amount of work performed in turning the generator and generating electricity. One can readily imagine the complexity of attempting to programme this whole problem at once. Instead, the building block approach is used and a separate programme is written for each piece of equipment involved, some forty programmes in all.

These programmes are arranged in the computer memory, which may be a magnetic drum as shown in Fig. 9, and along with each programme space is provided for entering the data concerning each device of that kind in the system. There is also space to indicate to what other devices each one is connected. Altogether, this completely describes the steam system to the computer. The calculations made by the computer then treat each device singly, assuming that the others are operating at the value at which they were last calculated. It proceeds through the system in much the same way as the steam flows through the system, calculating one device after another, and

determining in each case the new conditions for each element. After it has proceeded around this loop several times, if the problem has been properly planned, the solution converges to a steady-state value which is the correct heat balance for the system. Thus this tremendous problem has been solved without the engineer ever having understood all the detailed intricacies of



HEAT BALANCE CALCULATION ON 704 DIGITAL COMPUTER

Fig. 9.

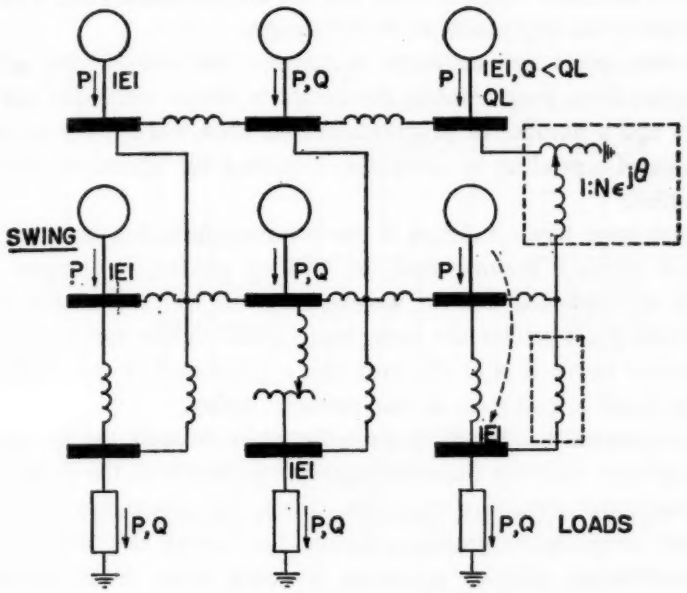
the system but having only a general idea of the interrelations and having programmed in detail each individual device and the interconnexions of the system. This is the building block approach.

This same approach is used for problems in widely different fields. For example, Fig. 10 shows an electrical problem of power systems. In this case the building blocks are busses (shown in solid black); they may be generating station busses, transmission busses, or load busses of different kinds. Altogether, in this figure nine different building blocks are shown and a separate programme is written for each.

When electrical equipment—for example, generators, transformers and circuit breakers—is applied in an electric power system, the principal problem to be solved is that of the load flow. Just as in a steam system it is necessary to determine where all the steam flows, so in an electrical system it is necessary to determine where all the electricity flows in order to determine the required sizes of all the devices. For any particular power system, it is first necessary to identify each bus as one of the available building blocks, and to provide the data to the computer on this basis. Then the com-

puter proceeds through the system, bus by bus, calculating all the currents and voltages and recalculating them until no further change is noted. This is then the solution for the system.

Fig. 10 LOAD FLOW BUILDING BLOCKS



Programmes of this type are so general that they may be thought of as languages. For example, the programme cited for steam heat balance will calculate the heat balance of any system composed of any combination or arrangement of those 40 different types of building blocks and including any number of each type. The same is true of the programme for the electrical system, which will calculate the load flow for a power system of almost any size using the building blocks in any possible combination. This ability to describe the system to the computer by merely identifying the building blocks may be considered as a language. A language oriented to the steam engineer nows exists for describing a steam system to the computer, and there is also a language for describing an electrical power system to the computer, a language that is oriented to the electrical engineer. In each case a programme has been written which generates a full-fledged computer language programme and proceeds to solve the particular problem. Similarly, languages have been written for machine translation of human languages, for point mechanics, for machine tool control, and for many other specific applications of computers.

COMPUTERS IN MANUFACTURING

Let us turn our attention for a moment to the use of computers in manufacturing. When the large-scale computer made its appearance it was first applied to the solution of certain of the difficult scientific problems which arose in design, and almost all engineering programmes up to four years ago were of this kind. Most of them still are not complete designs but simply the solution to the key problems in the design.

About four years ago, engineers recognized that considerable advantages would accrue from programming the complete design technique and already two years ago a number of programmes had been worked out for the complete design of a product by computer, including the optimizing designs described earlier.

The next stage beyond design is the complete handling of an order from the time it enters a factory until the finished product is shipped. Already this is an accomplished fact for warehouse items which need not be manufactured and great strides are now being made in the similar handling of manufactured items. It is in this area that a great deal of the computer programming work is going on in our factories today.

The programme for handling an order right through the factory should not be confused with the mass business programmes of the factory dealing with pay-roll, the control of materials, production scheduling, and accounting. These computer programmes handle the factory information in mass; but it nevertheless remains necessary for each order to be handled right through, drawing information from the mass programme and feeding information back to it regarding the particular order.

COMPUTERS IN CONTROL

To illustrate the use of computers in control systems, I have chosen a few examples, the first of which (Fig. 11) is a college experiment conducted at Case University in Cleveland, with the co-operation of Westinghouse and several other companies. This illustrates the principal somewhat better than actual installations because it is simpler. The process studied was the hydrogenation of cottonseed oil and the objective of the computer was simply to increase the production and make a better product. The complete equations of the process were very complex but they were linearized in the region of operation. The computer was divided into two parts. The first part determines the constants in the equations that form a mathematical model of the system. The second part uses this model in determining the optimum control.

To provide data for the first part of the computer, the process was sampled every five minutes and the computer then solved the simultaneous equations necessary to make the mathematical model fit the last five samples.

This mathematical model was then used in the second part of the computer, varying the controllable parameters and integrating to the end of the process, to determine the best setting of the controls to complete the process as quickly as possible with the proper product mix. The result of the experiment was a 25 per cent reduction in production time and 10 to 1 improvement in the control of the end product. The success of this experiment established a milestone in the application of computers to process controls of all kinds.

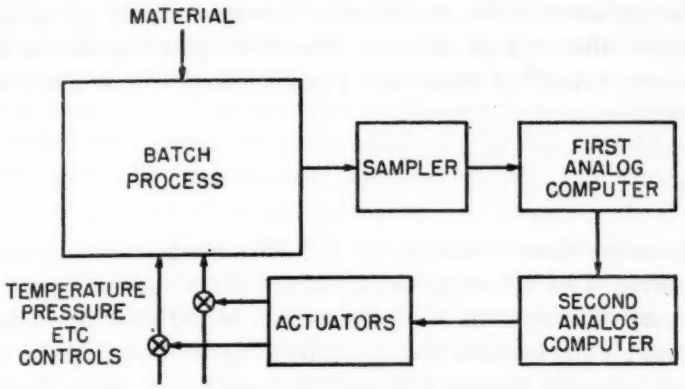


Fig. 11. Computer Control of Hydrogenation Process.

Another type of computer control is the Prodac, a programmed digital automatic control for a steel blooming mill. The complete programme for rolling a particular bloom into plate is supplied to the computer control in the form of a deck of punched cards. This then controls automatically the reversing of the mill as the white-hot bloom moves back and forth, under the rolls, the screwdown of the rolls being quickly adjusted between each pass, the side guide adjustments being made automatically, and the turnover manipulators controlled. The mill can be quickly changed to the settings for the next order of steel simply by inserting another deck of cards without any loss of valuable mill time—the operation being completely reproduceable for another order of the same type to be rolled several months later. Computer controls for such applications are being developed so as to make the calculations based on the measurements during operation allowing for variations in temperature or composition and have already advanced far beyond the stage we have described.

Another computer is used for the economic dispatching of power to an electric power system. The operators of an electric power system are continually faced with the problem of determining the most economical loading of each of the generating stations and machines. This depends on the fuel costs and efficiencies of the various machines, but also on the transmission losses. The resulting system of equations are solved by this computer which then either indicates to the operators or directly controls the generation to the optimum value on each machine. This is an optimizing control and most such computers installed more than pay for themselves in about one year of operation.

BUSINESS

The business problems of inventories and accounting were mentioned under manufacturing. Two seemingly contradictory trends are taking place in the business uses of computers. At first the giant brains were developed somewhat separately for business and scientific purposes based on the view that business calculations usually required little calculation but a great deal of input and output information, whereas scientific calculation was just the opposite, requiring large calculations with very little input or output. However, it has been found that the input-output for scientific problems requires at least half the computing time unless magnetic-tape input and output is used. This has led to the use of magnetic tape for scientific problems as well as for business problems. Also, programmers of both types of problems have found that a large high-speed memory is extremely advantageous. Therefore, the trend is to build computers suitable for either type of problem, so that a single computer can be used in a large factory and supply both needs. However, for many businesses having a very active inventory of from 50,000 to some hundreds of thousands of items, completely random access is highly desirable, so that much smaller and lower-cost computers with large random access are now being widely applied for these purposes, each computer being tied up completely on a single operation for which it is intended. These applications are all characterized by a large inventory of information to which access must be had frequently. In general, it works out that if the information must be searched over at least once a day on magnetic tape, it is about as economical to have it in a random access store of one second or less access time where it is kept available continuously.

One of the most interesting and promising areas of computer application in business is that of operations research, production planning and scheduling, and business decisions. Some very useful papers of the Unesco Conference on Information Processing dealt with these operations research techniques. On a computer, a business or a large section of a business can be simulated in order to study alternatives in the operation of the business,

an operation which is totally impossible without a computer. These are the difficult problems, and often those which might most profitably be solved, on which much effort will be spent over the next several years.

CONCLUSIONS

In this brief article it has not been possible to mention vast areas of the use of computers—in government, in research, in the translation of languages, the retrieval of information, and other fields. We hope nevertheless that we have in some measure succeeded in illustrating the tremendous scope and the revolutionary nature of this development. In less than fifteen years it has grown from a few tiny scientific applications to a point where it now reaches into every avenue of human endeavour. The powers of the computer, the heart of the development, are still increasing at a rate estimated to be some 10 to 1 per year, and there is little evidence of their reaching a limit. The fact that applications lag at present many years behind the available technology offers a tremendous challenge to the engineers and scientists of the world. Far from depriving man from employment, the use of this tremendous technology will lighten man's mental burdens and elevate him to a new world of greater effectiveness and greater opportunity, just as has the harnessing of the forces of nature to provide his physical power and transportation.

MACHINE TRANSLATION AND THE HUMAN BEING

by
D. Y. PANOV

D. Y. Panov, Doctor of Science, Professor at the Moscow Institute of Physics and Technology, was the organizer and first Director of the All-Union Institute for Scientific and Technological Documentation of the Union of Soviet Socialist Republics. In 1954 he set up, in the institute, a working party for reasearch into machine translation—the first of its kind in the USSR.

The research director of a large industrial corporation recently stated that, provided the cost of research into a particular subject was not prohibitive, he preferred to begin such research from the beginning, without looking up any results which might already have been achieved and published—because inquiries of that kind took up so much time and it ultimately cost more to solve problems on the basis of already-published results than to tackle them afresh. This statement may be an exaggeration; however, reference to the results of scientific research does constitute, in our day, a serious issue.

The number of scientific papers published throughout the world, and of publications relating to each of the separate branches of science, increases from year to year. In chemistry and the related fields, more than 50,000 books, articles, patent specifications, etc., are issued yearly. In other words, about 150 new publications dealing with chemistry make their appearance every day of the year, holidays included. Thus, to extract information required from this flood of reported results is obviously an overwhelming task. The difficulty is increased by the fact that the new papers are published in every language in the world, including a regular and rapidly increasing proportion in languages described as 'difficult of access'. The list of journals published in Russian, which many people still regard as a 'difficult' language, contains over 1,000 titles; about 700 journals are published in Japanese, and so on. Scientists would like to have access to all this material; but this is frequently impossible, even if they can secure the actual article, book or specification, because they do not understand the language in which it is written. And here the problem of translation becomes acute. As we know, translation is a costly and time-consuming business; and, what is even more important, there are not so very many competent translators in the world. A good scientific translation is extremely hard to obtain, because the translator is seldom familiar with the branch of science to which the article relates. When the article to be translated is a general one, the translator

often feels so sure of himself that he produces an entirely subjective rendering.

Moreover, scientific and technical material is not the only kind of matter that needs to be translated. General material is required as well—extracts from newspapers and general reviews, information given over the wireless, etc. Translations are needed in correspondence, whether by letter or telegraph; and, of course, they are essential to the successful functioning of international organizations. It is no exaggeration to say that the problem of translation is one of the most serious with which present-day mankind has to deal, since upon its successful solution depends our awareness of what is going on in the world, our understanding of other nations, the dissemination of knowledge and the possibility of establishing international co-operation in every sphere of intellectual activity.

MACHINE TRANSLATION

A very attractive way of dealing with the problem is presented by the use, for translation purposes, of electronic machines, similar to those employed for mathematical computation. At first sight the idea of translating from one language into another by mechanical means seems almost fantastic; but a minimum of reflection on the subject will show us that there is nothing incredible about it. A language is, after all, merely a specific system for giving material expression to ideas and concepts. The different meanings of words, together with any variants of meaning, are given in dictionaries and grammars. There are no unknown words in any language, except those that are deliberately invented, and no concept that cannot be conveyed by the order in which the words are placed and by the connexion between the different words. Otherwise, a language would not fulfil its main function, which is to provide a means of human intercourse; it would be impossible to use it as a means of communicating one's thoughts to another person. But this, as we all know from experience, is being done all the time; in conversation and in reading we are able to discern extremely delicate shades of meaning, expressed in words. So obviously there must be some material means of establishing a precise, objective distinction between them. The essential thing, for this purpose, is that the distinction shall be drawn according to definitely established rules, and this is exactly what machine translation renders possible.

To mechanize the search for a required word in the dictionary is easy enough. All that is necessary is to write it out in some selected code (such as the Baudot telegraphic code) and then, with the help of the machine, to compare this word with all the other words in the dictionary. These other words, of course, must be written in the same code in the 'memory' mechanism of the machine. Difficulty arises when a dictionary does not give a

single, uniform meaning for the word—in the case of polysemantic words, for instance. But even here the question can be solved by properly analysing a few aspects of form.

A translator, when choosing between the possible meanings of any particular word he encounters, is usually guided by the context; that is, he considers the word together with whatever other words accompany it. For instance, the English adjective 'sweet' may have at least six different meanings in Russian. The choice between these meanings will be determined by the noun that this adjective qualifies. The Russian translation will use different adjectives to render 'sweet apple', 'sweet song', 'sweet words', 'sweet girl', and so on. But the right term can be selected by purely mechanical means, with the same formal symbols being used. It is entirely a matter of correct programming—of showing clearly and precisely what has to be done, and in what order. If this is established, the work can be carried out by a machine. Problems of this kind are more complicated than the calculation of definite integrals or the solution of differential equations with which machines frequently have to deal; but the machine is quite able to cope with them.

THE HISTORY OF MACHINE TRANSLATION

The first attempt to mechanize translation appears to have been made in 1933 by a Soviet citizen, P. P. Troyansky. He proposed the construction of a machine which would select and print words in the process of translation from one language into another or into several different languages simultaneously.[1] A certificate of authorship was issued to P. P. Troyansky for this invention, but it was not actually followed by the construction of the machine. This is understandable, since the necessary mechanisms did not exist at the time. Troyansky suggested that the translation should be made with the assistance of two editors. One would be acquainted only with the language from which the translation was to be made (the 'source language') and would supply the text with the aid of 'logical analysis symbols' indicating the sentence-construction, the parts of speech and their form. The other, knowing only the language into which the translation was to be made (the 'target language'), would take the text delivered by the machine and convert it into current phraseology in that language, with the help of the accompanying logical analysis symbols. It is interesting to note, however, that even at this early stage P. P. Troyansky wrote that '. . . the process of logical analysis could itself be mechanized, by means of a machine specially constructed for the purpose . . .'.

American and English scientists began to concern themselves with the problem of machine translation in 1946. Initially they, too, considered that

1. *P. P. Troyansky's Translating Machine*, published by the USSR Academy of Sciences, Moscow, 1959.

The process would involve editing, at least at the final stage. In 1954 the International Business Machines Corporation, in co-operation with the University of Georgia, carried out a celebrated experiment in translation from Russian into English, using the IBM-701 electronic computer. Although only very simple Russian phrases were chosen for translation (see examples in Table 1) and the mechanized vocabulary was very small (250 Russian words), this experiment proved that machine translation, without the need of editing, was in principle feasible.

TABLE 1

Качество угля определяется калорийностью.	The quality of coal is determined by calory content.
Крахмал вырабатывается механическим путем из картофеля.	Starch is produced by mechanical methods from potatoes.
Обработка повышает качество нефти.	Processing improves the quality of crude oil.

At the end of 1955 the USSR Academy of Sciences, using a BESM electronic computer, made the first experiments in the translation of entire passages of a scientific text from English into Russian. For these experiments a vocabulary of approximately 1,000 English words was used, and here again no editing was required (see examples in Table 2).

TABLE 2

In problems of this type numerical methods become a necessity due to absence of other methods for getting the requisite information out of the differential equations.	В задачах этого типа численные методы становятся необходимостью, обусловленной отсутствием других методов для получения необходимого сведения из дифференциальных уравнений.
Even in cases where explicit or implicit solutions are known, it is sometimes easier to obtain a numerical solution than attempt to calculate numerical values from the known solution.	Даже в случаях, где явные или неявные решения известны, иногда более легко получить численное решение, чем пытаться вычислить численные значения из известного решения.

Research in the field of automatic translation is at present proceeding in the USSR, England, France, Japan, China, Sweden and other countries. Algorithms of translation from English into Russian and Japanese, from Russian into English, from French and German into Russian and English, from Chinese, Japanese, Magyar and a number of other languages into Russian, and so on, are being studied.

TWO ASPECTS OF THE PROBLEM

It has now been demonstrated that electronic computers can be used to obtain translations, from one language into another, which do not require editing. But the first translations, carried out with universal computers, were clearly unsatisfactory. A computer used for translation purposes is working much below capacity, for translation enlists only a few of its resources. For this reason, if mechanical translation is to become worth while in practice it will be necessary to construct special electronic translation machines, of the same type as the computers but of rather simpler structure in the sections performing arithmetical operations. And here there arises a dilemma on the solution of which the further development of work on machine translation will depend.

The answer to the problem of machine translation depends on two factors —the actual machine, and the algorism of translation, which is indispensable for preparing the machine's work programme. Opinions may differ as to the relative importance of these two aspects of the matter, some people being inclined to give priority to the machine and others to the algorism. To put the question more simply, we may say that holders of the former view would be in favour of constructing a more powerful machine and using a simpler algorism, while holders of the latter view would prefer to work out a more elaborate and flexible algorism and to use a less powerful machine. The matter must be considered, to a certain extent, in relation to the problem of the vocabulary for automatic translation. There may be one large dictionary, aiming at universality and suitable for the translation of texts dealing with a variety of subjects; or there may be a series of restricted dictionaries, known as 'micro-dictionaries', each of which is intended for the translation of texts dealing with only one field (physics, radio-technology, mathematics, etc.); when a text is to be translated, the appropriate micro-dictionary must be inserted in the machine. A translation machine with an extensive vocabulary must have a very capacious memorizing apparatus and the time it takes to select words ('access time') must be brief. Such machines would have to be specially designed, for apparatus as adaptable and reliable as this does not yet exist. On the other hand, the existing types of memorizing mechanism are quite sufficient for dealing with micro-dictionaries, and the method of inserting them, when required, presents no difficulty. In short, if the trend of development is to be in the direction of powerful, highly-specialized machines, working with a comparatively simple algorism, they will have to be specially designed; if the trend is to be towards a more complex algorism, it will be possible to construct machines on the basis already established in the field of computers. We consider the latter course to be the better, since it would yield practical results more rapidly. The possibility of working out the necessary algorisms in this event is

clearly demonstrated by the example of the English-Russian algorism of automatic translation prepared in the USSR and tested with satisfactory results.

AUTOMATIC AND MACHINE TRANSLATION

The terms automatic translation and machine translation are often employed as though they were synonymous. But this is not quite the case. The term 'automatic translation' has a wider meaning than the term 'machine translation', since every machine translation is automatic, whereas an automatic translation is not necessarily made by a machine. It may be made by human beings, but in obedience to special rules leaving no initiative to the translator, who is thus obliged to work in a completely automatic manner. In both instances there has to be an algorism of translation; but for automatic translations, carried out by human agency, the algorism may be written out in ordinary language, whereas for machine translations it must be adjusted to the machine's work programmes written in the appropriate code. This algorism, without which there could be no machine translation, produces a further result: it establishes simple and objective rules for translation which, if scrupulously followed, ensure that the translations will be accurate and linguistically correct. It establishes its own translation norms, and produces standardized translation.

This seems to us to be of considerable significance. We have no doubt that, within a few years, machine translation will have great successes to its credit and that the machines will to a large extent relieve human beings of the task of translation. But it would perhaps be over-optimistic to suggest that even 15 or 20 years hence the need for 'hand-made' translations will have ceased to exist—particularly in view of the spread of scientific knowledge and education among the populations of a number of countries which, until recently, had no possibility of acquiring them. The population of those countries runs into hundreds of millions, and even on the most optimistic basis it is difficult to suppose that all the translation problems arising in this connexion could be solved by machines alone. But if human translators have to be brought in, the existence of rules of automatic translation will do much to ease their task and deliver them from the worst calamity that can befall a translator—the necessity of editing and correcting bad translations, which reduce the qualified translator to despair and take up more of his time than if he had done the work himself from the start. The algorism of an automatic translation, once it has been prepared, can be either employed for the preparation of a work programme for a translation machine, or published in book form, as a guide for inexperienced translators or even for those who do not know the language concerned, and as a work of reference for experienced translators.

THE ALGORISM OF AUTOMATIC TRANSLATION

The algorism which has so far been most thoroughly elaborated seems to be that used for translations from English into Russian. It consists of the following parts: (a) a scheme for looking up words and determining the meaning of words with more than one significance; (b) a scheme for analysing English sentences; (c) a scheme for constructing Russian sentences.

The last part of the algorism—the scheme for constructing Russian sentences—is the least interesting, since it works without difficulty in accordance with the established rules of grammar. For the purpose of automatic translation without a machine it is quite superfluous, since anybody acquainted with the Russian language can rapidly put together the required Russian sentence without the help of a system, from the data worked out by parts (a) and (b) of the algorism.

The dictionary work is covered by part (a) of the algorism. As already stated, we consider that the correct method to adopt is that of the micro-dictionaries. Experience has shown that a vocabulary of about 2,500 English words is sufficient for the translation of any text dealing with applied mathematics. This gives grounds for assuming that a dictionary comprising between 2,500 and 3,000 words is amply sufficient, and that quite a wide range of subjects can be covered by dictionaries of that size. The use of micro-dictionaries not only reduces the demands on the machine's memorizing mechanism; it also simplifies the analysis of the sentence to be translated, because in micro-dictionaries synonyms are cut down to a minimum, as also are the number of words with more than one meaning. In automatic translation, as has already been explained, the precise significance of a word with more than one meaning is ascertained by analysing the context. This process deserves to be looked into, for it explains the resemblance between the methods of solving problems of automatic translation and those applying to a number of entirely different logical problems. It supplies a typical example of dichotomy, very similar to the process adopted when determining the species of a flower or butterfly in botanical or entomological classification.

Table 3 illustrates the simplified system of selecting a meaning, applied to the polysemantic words 'many' and 'much'.

The symbols used in this scheme are as follows: the formula A (B, C) signifies that if the question asked at A brings an affirmative reply, we should proceed to check point B; whereas if the reply to A is negative, we proceed to check point C. The symbol A (0) indicates a final answer, which renders further checking superfluous. We will now see how the system works in specific instances.

1. 'The subject would have been *much* better standardized.'

In line 1 of the scheme in Table 3, we have to check whether the word

TABLE 3

1(2, 3)	Check the immediately preceding word for 'how'.
2(0)	Skol'ko (numeral, invariable).
3(4, 5)	Check the immediately preceding word for 'as'.
4(0)	Stol'ko ze (numeral, variable).
5(7, 9)	Check given word for 'much'.
6(0)	Not to be translated (adverb).
7(1, 11)	Check the immediately preceding word for 'very'.
8(0)	Mnogij (adjective, hard stem, with a sibilant).
9(8, 12)	Check preceding word for preposition, and following word for noun.
10(0)	Mnogo (adverb).
11(12, 10)	Check following word for noun.
12(0)	Mnogo (numeral, variable).

'how' does not precede the word 'many' or 'much'. That word does not appear in the sentence, so we have a negative reply and must go on to line 3. Here we have to check whether the preceding word is 'as'. Again the answer is in the negative, so we proceed to line 5. Here we check whether the word is 'much'. The answer is in the affirmative, and we go on to line 7. We look at the preceding word and find that it is not the word 'very'. We proceed to line 11 and look at the following word. In this instance that word is 'better', i.e., is not a noun; so we must move on to line 12, which gives us the final translation, 'mnogo' (adverb).

2. 'This is most useful, and for *many* reasons.'
We proceed in the following successive stages: 1—the answer is 'no', so we proceed to 3; 3—the answer is 'no' so we proceed to 5; 5—the answer is 'no', so we proceed to 9; 9—the answer is 'yes', so we proceed to 8; 8—gives the translation, 'mnogij' (adjective).
3. 'There should be as *many* equations as there are unknown quantities.'
1—the answer is 'no', so we proceed to 3; 3—the answer is 'yes', so we proceed to 4; 4—gives the translation, 'stol'ko ze' (numerative).

These examples show clearly how the meaning of a polysemantic word is determined in the process of automatic translation, and also reveal the restrictions that the procedure naturally entails. If a word has too many meanings, the scheme becomes extremely complex, and its use as a working method would cause waste of time owing to the necessity of analysing a great many variants, only one of which would actually be needed. For this very reason it is undesirable to try to set up large dictionaries embracing a wide range of subjects. It is preferable to have highly specialized micro-dictionaries with a strictly limited number of polysemantic words.

The most important function of an algorism of automatic translation is the analysis of the sentence which is to be translated. For it is this analysis which yields all the grammatical features of the individual words covered by it, and makes a correct translation possible. And from the linguistic

standpoint the most important and interesting result achieved in the sphere of automatic translation is I. K. Belskaya's discovery that this part of the algorism can be applied universally, irrespective of the texts for which the algorism is actually intended. Belskaya demonstrated that the systems of analysing English, based on a study of texts in applied mathematics, can also be successfully applied to literary translation as well. For example, the scheme of analysis based on E. Milne's thesis 'Numerical solution of differential equations', Hardy's 'Divergent series', and so forth is perfectly suitable for the analysis of writings by Dickens, Galsworthy or Aldridge.[1] It will only provide a standardized translation, with no pretensions to 'artistic' merit; but this fact shows very convincingly that, in solving the problem of working out an algorism of automatic translation, the linguist should combine in a clear and accurate system the chief elements in the language which form the basis of all writing, whether scientific or literary.

AUTOMATIC TRANSLATION AND LOGICAL PROPOSITIONS

Among the problems now being solved by computers, increasing importance is attributed to what are known as logical propositions. In these propositions it is sought to discover one solution or another depending on the fulfilment or nonfulfilment of certain conditions. The planning of production, the problem of making the best use of transport, and so forth, come under this heading. The problem of automatic translation falls within the same category. Earlier in this article we gave an example of one of the schemes used for the solution of that problem. And that scheme is a typical example of the handling of logical propositions. In addition to the linguistic aspect, which we have been considering up to now, the problem of automatic translation also has a mathematical aspect, connected with the preparation of a work programme for the machine, based on an algorism which has to be evolved. This question can be settled without necessarily considering the specific linguistic content of the various sections of the algorism, and we are then confronted with the need to devise means of solving extremely complex propositions in logic. Each successful construction of a programme for automatic translation will undoubtedly be equivalent to a success in the development of programmes for dealing with other logical problems.

CONCLUSION

Machine translation from one language into another is one of the most interesting scientific and practical problems confronting mankind at the

1. I. K. Belskaya, 'Machine translation methods and their application to [the] Anglo-Russian scheme', paper read at the International Conference on Information Processing, Unesco, 1959.

present time. Its solution is of interest to all nations and to all international organizations, particularly those like Unesco. International co-operation between scientists and manufacturers would be of great value in dealing with this matter, and it might be worth considering the possibility of setting up a special international scientific organization for that purpose.

MACHINE TRANSLATION OF LANGUAGES: RESEARCH AND ORGANIZATIONAL PROBLEMS

by

E. DELAVENAY

E. Delavenay, Chief of the Documents and Publications Service of Unesco since 1950, was formerly Chief of the Official Records Division of the United Nations. He has been a teacher of English language and literature, and has taught French literature in English universities. His present functions as well as his personal interest have led him to take up the study of the applications of computer techniques to the problems of linguistics and international cultural relations, with which he has been concerned throughout his career.

INTRODUCTORY REMARKS

It was not so long ago that some of the most intelligent of men were still dubious about whether translation itself was really possible—a paradox that reminds us of Zeno of Elea denying the possibility of motion. Yet the perfectionists are given the lie by the completely bilingual 5-year-old child who can translate accurately into French what his kindergarten teacher tells him in English; and who may thus be said to have proved the reality of movement by walking!

We may therefore take it as axiomatic, or as a patent fact, that translation is possible, while provisionally granting the existence of a special preserve in which perfect translation is no more conceivable than an absolutely perfect performance of a piece of music or an ideal copy of a painting by a great master. This special field is that of literary translation, as against purely informative or 'indicative' translation or—to put the matter briefly and at the risk of rather over-simplifying a side-issue—as against scientific and technical translation.

Recently, the question has been mooted of the possibility of mechanical translation, that is, translation done entirely, or almost entirely, by a machine. This question, too, has been answered in the affirmative, and there are some optimistic spirits who have such faith in their analytical methods that they have even broken into the hallowed preserve described a moment ago. Miss Belskaya and Professor Panov vigorously maintain—and use excellent arguments to back their views—that machine translation of poetry will be feasible within the near future.

Let us assume, then, that machine translation is already possible for many types of work, and that further research and organization are all that is

needed. Let us also admit, provisionally, that it has certain limitations, just like any other human undertaking—limits to its scope of application and to its full automation.

FROM MAN TO ROBOT: PARALLELISM OF FUNCTIONS AND ORGANS

Professor Panov was, I think, the first to expound the principles of automation in translation work by drawing a very simple parallel between the work of the human translator and of the machine which can replace him. That parallel, in fact, cannot be bettered as an aid in trying to define the organs and operations of the machine designed to achieve the end in view: the production of a typewritten translation.

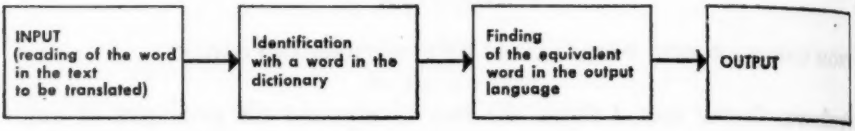
Setting a machine to do the work of a translator means automating a highly complex process, the main stages of which are as follows.

When I translate a written sentence from English into French:

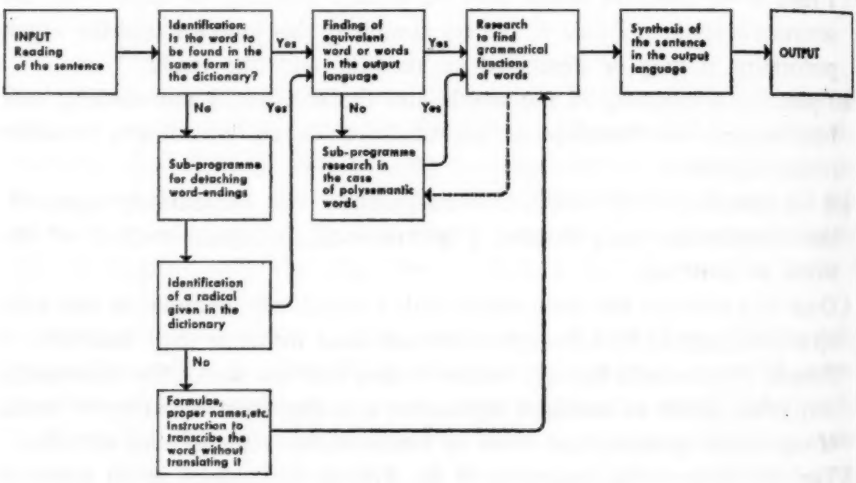
1. I read it—a function which may be termed 'reading' or 'input'. (If the sentence is spoken, I hear it, but for simplicity this function and the organ performing it may be described as 'reading' and 'the reader'.)
2. I perceive a meaning in the words and the sentence I am reading; this function may be described as 'identification' of the constituent elements of the sentence.
3. If the meaning of the words or word groups is not immediately apparent, this identification may involve a grammatical or logical analysis of the words or sentence.
4. Once that analysis has been completed, I search my memory or else look up a dictionary to find the equivalents of these words or their functions in French. A moment's thought makes it clear that the search for equivalents may relate either to words of equivalent meaning (semantics) or to words of equivalent grammatical value or function (morphology and syntax).
5. The next stage is the grouping of the French equivalents in an order in harmony with French usage and grammatical rules. This series of operations (establishment of word order and grammatical agreement) may be described as 'synthesis'.
6. Lastly comes the psycho-muscular operation of giving vocal expression to the translation of my sentence or inscribing it on a sheet of paper. This last operation may be termed 'writing' or 'output'.

Reading or input, identification of the input material by comparison with words stored in the memory, analysis when the words are not simple or do not coincide exactly with those stored in the memory, the search for equivalent words in the second language and their registration in a memory cell, combination or synthesis of these words in order to form a French sentence, writing or output—these, more or less, are the stages of my work as a translator which the machine will also have to reproduce.

When I am faced with an extremely simple translation problem, my work is in accordance with the following pattern.


INPUT
(reading of the word in the text to be translated)
Identification with a word in the dictionary
Finding of the equivalent word in the output language
OUTPUT

If I wish to translate a single word, such as 'paper', it is easy to find the simple equivalent, *papier,* assuming that this word has no other meaning. The same applies to simple groups of words, such as 'the little horse', *le petit cheval;* 'the cat eats the mouse', *le chat mange la souris.* But problems are involved even with such words as 'le', 'petit', 'cat' and particularly with the inflected word 'eats', and the programme to be carried out is then more complex (see figure below).


INPUT
Reading of the sentence
Identification: Is the word to be found in the same form in the dictionary?
Yes
No
Yes
Finding of equivalent word or words in the output language
Yes
No
Yes
Research to find grammatical functions of words
Synthesis of the sentence in the output language
OUTPUT
Sub-programme for detaching word-endings
Sub-programme research in the case of polysemantic words
Identification of a radical given in the dictionary
No
Formulae, proper names, etc. Instruction to transcribe the word without translating it

Once the meanings and values of words and their combinations have been identified it is relatively easy, even for a machine, to translate the sentence.

But how is this work to be done? And what organs will be used?

Human beings read with their eyes, the visual impressions received being transmitted by the nervous system.

Identification is effected by means of nerve centres, assisted by the eyes where the dictionary has to be consulted.

Analysis is carried out externally, with pencil and paper, and internally, in the nerve or brain centres, by processes which have not yet been subjected to searching investigation.

Equivalent words are found either in the human memory or in the dictionary (which may be regarded as an auxiliary memory).

Synthesis is carried out either in the memory in the case of a simple sentence, or on paper in the case of more complicated sentences.

The final result (output) is recorded by hand, in writing or in typescript.

The 'organs' used during this process are also to be found in a translating machine.

The reading is done by means of punched cards or tapes or, preferably, by magnetic tapes; an even better method, which will be feasible once certain technical problems have been solved, will be the use of photo-electric cells or some other photoscopic process.

The identification is done by comparing signals representing a word fed into the machine with signals stored in the memory-dictionary. This is a simple subtraction operation if the signals are binary digits. Other methods apart from arithmetical subtraction are conceivable. In every case, the operations are carried out in circuits which are connected up in much the same way as the neurones of the nervous system.

The analysis is conducted by means of what are termed sub-programmes. In this case, the parallel with performances by human organs can still be drawn, though with difficulty, provided we consider the matter in terms of actions rather than static conditions.

Let us take an example. The word 'loved' is fed into the machine. The dictionary does not show this word, which must therefore be looked up in a different form. A sub-programme for detaching word endings now comes into play. At a given stage, the machine will be asked the question 'Are the last two letters of the word ed?' and will give an affirmative reply. It will then be immediately instructed to look up the rest of the word (i.e., lov-) in the dictionary, where the French translation will be found. Once the word-endinged has been identified as indicating the past tense in English, an appropriate symbol will be entered in a memory or register (the equivalent of the sheet of paper used by a schoolboy for making his translation notes). The sub-programme for detaching endings will thus make it possible to identify simultaneously, or almost so: (a) the ending of the word to be translated; (b) its radical; (c) the ordinary meaning of this radical in French; (d) its grammatical function (e.g., verb); (e) its tense (in the case of a verb).

There are of course other features relating to the word which still have to be ascertained (such as its subject, and whether the latter is singular or plural, and what type of complement the word governs, etc.). Other sub-programmes for dealing with the rest of the words in the sentence will make it possible to store in the machine's memory all the data necessary for sentence construction in French.

Let us now revert to the other operations involved in an over-all programme of machine translation. The last one mentioned was identification, which in this case implies a series of analyses. (As we have already seen, identification is often accompanied by analysis and a search for equivalent words.)

These operations are carried out in the electronic circuits of the machine through the agency of successive sub-programmes. They may be described briefly as the putting of a series of questions to which the answer will be either yes or no, each reply leading either to the registration in a memory organ of a result identified by the mechanism, or to a further search, which may culminate in turn in either a definite result or a renewed search. The sub-programmes may thus be compared to a succession of switches on a railway line, each switch either bringing the train to a standstill (recording of a result) or directing it to a further destination (continuation of the search). The same applies to synthesis, of which we shall now try to obtain a clear idea. Let us revert to our simple sentence, adding another word so that it becomes: 'The cat eats the white mouse'. Let us assume that the machine has identified each of the words in the sentence as follows: (a) definite article (gender to be determined); (b) noun rendered in French by the word *chat,* masc. sg.; (c) verb meaning *manger,* third person, present indicative; (d) definite article (gender to be determined); (e) adjective meaning *blanc* which has to agree with the following noun; (f) noun meaning *souris,* fem. sg.; (g) noun and adjective to be placed in correct order.

To write the sentence *Le chat mange la souris blanche*—that is to make a synthesis of the words in the sentence (or in school terms) to produce a French translation—the machine will have to arrange the French words found in the dictionary in the correct order, with nouns and adjectives agreeing in number and gender and verbs in the correct tense. To do this it will consult tables. There will be tables of definite articles giving the word *le* before *chat* because the latter is masculine, and *la* before the feminine word *souris;* a table of word order, which will make it say *souris blanche* and not *blanche souris,* and so on. This, very briefly, is the process of synthesis. To avoid elaboration, all that need be said here is that this process of consulting tables is just as feasible for synthesis as it is for analysis, and can be rendered entirely automatic, as it is in computation. The real problem is to build up valid tables which are not unduly complicated.

So far as output goes, suffice it to say that the signals registered in some kind of electronic memory will be able, when required, to work a typewriter, just as my hand writes or my mouth utters the sentence I have translated from English whenever I please.

For the sake of clarity, we have deliberately simplified our initial exposition. Simplicity of operation, however, obviously does not come about automatically, but is the fruit of prolonged and extensive effort. In the sequence of operations described above, the obvious aim of each of these operations is to reduce work to the simplest possible pattern, the identification of each separate word in the source language being immediately followed by its translation. But the twists and turns of human language, the social and cultural differences between peoples, the growth of languages

and the existence of linguistic structures that are at once fossilized and living stand in the way of achieving such simplification, and our efforts will depend upon an initial act of faith in the possibility of doing so.

BRIEF HISTORY OF MACHINE TRANSLATION RESEARCH

While the idea of making the consultation of bilingual dictionaries completely automatic may not have struck fifth-form boys obliged to look up the meaning of words in a Latin dictionary running to 800 or 1,000 pages as a practical proposition, it was bound to occur to present-day scientists similarly obliged to read the reports of their colleagues in an increasing number of unfamiliar languages. The pioneers in this field were Smirnov-Troyansky and Booth, who, in 1933 and 1946 respectively, attempted to mechanize the consultation of dictionaries with a view to the production of more or less acceptable translations in telegraphic style.

Appetites were whetted by these initial endeavours, and as early as 1952, when the first conference was convened by the Massachusetts Institute of Technology (MIT), with the help of the Rockefeller Foundation, the idea took shape of going much further and rendering the translation process fully mechanized, or almost so. We have reached a point today where theoretical discussion as to how far translation can be mechanized is almost academic: between the simple mechanical consultation of the electronic dictionary and the 99.5 or even 100 per cent automation of translation lies a whole range of possibilities, all of which can lay claim to theoretical and practical value.

The following dates are landmarks in the history of machine translation.

1946: First discussion of the problem of mechanical dictionaries between Booth and Warren Weaver.

1946-1948: Work by Booth and Britten at Princeton on the automation of dictionaries.

1948: Richens suggests the automatic stripping of word-endings.

1949: Warren Weaver's historic memorandum gives a new impetus to research by drawing the attention of American scientists to certain aspects of language study capable of close analysis. In particular, Weaver induces the American sinologist, Reifler, to explore the system of signals either contained in the alphabet or in extra-alphabetical signs, visible or invisible but perceptible by the human reader and recognizable by a machine (e.g., spacing, position, punctuation, etc.).

1950: Reifler's first report. Work by Oswald and Fletcher on German.

1952: Definition by the first MIT conference of two short-range research targets.

1954: Experiment in machine translation by Dostert, and Garvin and Sheridan on an IBM-701. First issue of the periodical *Mechanical Translation* edited by William Locke and Victor Yngve.

1955: Publication of the first collection of essays on machine translation, edited by Locke and Booth.

1955: Beginning of Soviet research by the Panov and Lyapunov teams.

The scope of research has been greatly expanded since 1955, as witness the many centres now researching in the subject in the Soviet Union and the United States of America, to say nothing of others in the United Kingdom, Japan and Italy.

The spade-work done between 1946 and 1955 clearly showed that research by any team looking for quick, practical results was bound to follow a definite cycle or rather 'spiral' of development, and we have every reason to hope that by keeping to this course we shall progressively approach the fuller automation of the translation process.

As we have seen, then, that Booth starts from the idea of the mechanical dictionary look-up. The machine is conceived of as a means of helping the English or American scientist to read papers by his Russian or Chinese colleague by expeditiously providing him without any effort whatever on his part, with the meaning of each word in the order in which they occur in the original. The scientist is expected to make sense of this word-for-word translation because of his familiarity with the subject.

The obvious objection to this procedure is that the same word can have many different meanings, and it is met by giving every one of the meanings leaving the scientist to make his own choice among them.

We linguists and professional translators obsessed with the niceties of language and the difficulties of our art, greeted Booth's proposal with superior smiles. But we were wrong, and Booth and Richens were right, as the results obtained by Professor Oettinger's team at Harvard were to prove.

A mechanical dictionary could be self-sufficient provided words and ideas fitted exactly. But this is not the case; words represent concepts which on the whole are arbitrary and ill-defined. We only have to consider such French words as *et, ou, à, par, de, être, avoir, oser, penser, estimer, paraître, aimer, donner, façonner, amicalement, seulement, annonceur, pain, charcuterie, jardin, outil,* and word sequences such as *il n'est pas venu, de plus en plus, à qui mieux mieux, se faire du souci, mettre les bouchées doubles* to realize that neither simple words nor groups of words are units which are strictly comparable with each other as regards their meaning, grammatical or logical role in a sentence, structure or possible relationship with other words. Just try to translate these words into English or Russian and you will be faced with many problems in finding exact equivalents. The only thing to do is to study each word in relation to its immediate context and the meaning of the sequence or sentence as a whole.

In view of the impossibility of making individual words the basis of machine translation, the first thing that had to be done was to try to break up words which are not simple so as to mechanize the analysis of their

constituent elements, (e.g., features denoting gender, number, tense, person, etc.). Richens showed that tables of paradigms, such as we find in grammars, could be built into the storage part of the machine. It was thus possible to confine the dictionary to the stems or 'bases' of words, with references back and forth to the tables of endings.

This fairly easily disposed of the problem involved in the changing forms of a word: declensions in Latin, German or Russian, different forms to denote gender and number in French, and conjugations in all languages where the verb takes different forms according to person, tense and mood.

The attempt to deal with morphological inflections, incidentally, shed practical light on a phenomenon with which we were thoroughly familiar in theory. Inflections provide valuable information about the grammatical function of words in the sentence and hence, about the relationships between the verb, the subject, the object and the various complements of the verb. In so far as the machine has to record in a memory a full description of the special features of each word in the sentence, the analysis of word-endings affords a means of condensing and simplifying this description and the preliminary research.

But some languages are almost completely devoid of inflections. English has preserved only a few endings such as -s, -ing, -ed, -er, -est and -th from the highly inflected Anglo-Saxon system. Yet English, with its rich vocabulary and exceptional flexibility, is capable of expressing every possible shade of thought. Hence in mechanizing translation, a study had to be made of the rules of syntax and language structures in order to derive from the positions of words in a sentence information about their functions which could not be ascertained from the words themselves or from their endings. It was a case of eliciting information about the function and even the nature of a word by studying its context rather that of analysing the word and its constituent elements.

The point is that the same word in English may, without any change in form be a noun or a verb or an adjective, and its grammatical role may well be impossible to determine properly except from its position in relation to other words. The same, I believe, is true of Chinese.

This means that research in language structure—with particular reference to Jespersen's work on analytic syntax and Charles Fries's studies of the structure of English—has now been put to practical use; and although we should avoid trying to explain everything by structural linguistics, this line of development has yielded good results. When I read an English sentence such as 'He gave the child a swim-suit', I can often identify the verb and its behaviour in relation to its complements and even without fully understanding the sentence, deduce that the first noun (child) is in the dative, the second (suit) in the accusative and that 'swim' has an attributive or adjectival function in relation to 'suit'.

The machine itself is capable, with the help of appropriate sub-programmes, of examining the order or words and their interrelationship and of recording the results in the appropriate cell of a register or memory until it has deciphered all the grammatical constituents of the sentence.

The idea of the automatic dictionary has thus led naturally to that of similarly automatic grammars to supplement it. Nor is there any contradiction between the two ideas: analysis has simply resulted in an attempt to endow the machine with all the tools used by a translator.

And so we must be wary of suggesting misleading alternatives and becoming ensnared by preconceived theories of translation or language. There are in fact three complementary and inseparable stages in the analysis of texts: analysis of the meaning of words culminating in the automatic dictionary; analysis of the form of words, culminating in tables of inflections; analysis of the relationships between words, culminating in tables of linguistic structures for a given language, and perhaps in tables of equivalent structures for another language. But this third type of analysis will bear on aspects of word relationships which have hitherto been given scant attention.

These three stages, corresponding to the three basic aspects of any grammatical analysis of language are, I repeat, inseparable, for it is impossible to obtain full information about the meaning of a sentence or word sequence from a glossary or from morphological tables, or from tables of syntax or language structure taken singly. Where the human translator relies on what he calls intuition ('I think this sentence must mean such-and-such'), the machine can only follow its pre-established programmes blindly and try by means of the three stages of analysis to work out a word-for-word or sentence-for-sentence equivalent of the original text fed into it. Not only is it impossible for the machine to dispense with any stage of analysis; it must come back to each in turn to complete its investigations. To attempt an abstract definition of the order in which the machine has to cope with vocabulary, morphology and syntax, would be to approach one of the most important research problems from the wrong angle. The approach will obviously depend on the type of language to be translated, while the relative importance of the three elements in 'deciphering' the text will depend on the language's special features. Whereas, for Latin and Russian, the starting point might be grammatical analysis with particular stress on morphology, for English the most important thing might be to find the verb and see how it fits into its context. Structural analysis is therefore the proper approach for deciphering English. However, such analysis is impossible—for any sentence—without lexical information, which means that we come back to the inevitable and primary task of consulting the dictionary and the word-endings.

SOME PROBLEMS OF METHOD

Our historical survey has led us directly to the problem of the methods to be used in applied research. One question we were able to consider was whether priority should be given to any one of the recognizable stages in language analysis for the purpose of machine translation.

We must now turn our attention for a while to the importance of general theories as compared with applied research. In particular, we have to consider whether research is as yet sufficiently advanced for the formulation of a general theory of language suitable for practical application or whether it would not be preferable to put all our energies into developing our knowledge and understanding of the phenomena of language and translation, with particular reference to the formal definition of the mutual behaviour of certain categories of words which are ill-defined or not defined at all in the traditional grammars.

All science moves forward in stages: hypotheses are advanced and must be subjected to repeated factual checks. Applied linguistics has now reached the point where the most urgent task is the preparation of a comprehensive and systematic inventory of scientifically observed facts.

But scientific observation implies an attempt to classify material in accordance with criteria based on observed characteristics. Nothing is more dangerous than to begin with a classification accepted *a priori* or based on minor or secondary characteristics. One of the scientific necessities in applied linguistics today is a reappraisal of classifications in general (classification of parts of speech, for example) which delves beneath the old terminology based on Latin and Greek grammatical usage and adduces new word classes based on the behaviour of words in their immediate context.

What I have in mind is not so much the methods and theories of the advocates of the structural approach as the modified form of structural linguistics flowing more or less spontaneously from studies of context or word sequence with a view to machine translation.

We should take good care not to waste our time in pointless discussion of general theories of language, or thought, or any other aspect of the problems before us. Professor Panov, rightly anxious to see research yielding quick and concrete results, warned his readers against placing undue confidence in the virtue of more or less general theories, and the attitude he specified is one of scientific empiricism.

Machine translation, as a new subject, has naturally inclined people to rely on general theories or would-be universal explanations whereby the problems can be strikingly presented to the general or newspaper public in a nutshell. It is tempting, of course, to use a simple formula for explaining everything, especially in the absence of any obligation to see that a concrete programme is successfully carried through under practical conditions.

The theory of information (or rather the mathematical theory of communication), cybernetics, structural linguistics and mathematical linguistics are very general terms often cited in connexion with the problems of machine translation. Far be it from me to belittle their importance. Each of these disciplines has its part to play, spotlighting some particular aspect of the innumerable problems involved in translation from one language to another.

In the same way, while the linguist's strong sense of the individuality of language may often set him at loggerheads with the logician, that is not to say that their viewpoints are irreconciliable. The linguist's analysis is perforce closer to dynamic reality, while the logician's amounts rather to a systematization of thought emphasizing the underlying unity of our mental processes. The linguist stands in need of the logician, but he will look askance at the suggestion that language—that means of self-expression built up of the collective memories of a community and reflecting the often highly personal means of analysis of the speaker—can be reduced to a strictly logical formal system. Such a reduction might indeed hold good from the purely logical point of view but be totally useless for elaborating a practical translation programme.

The same is true when we come to the systematic application of mathematics to language. For the time being at least, whether we like it or not, electronic computers are designed to handle figures and figures only. It has been found possible to encode letters of the alphabet in digital form so that they can be fed into the machine, and this system has its advantages. The question is, will the situation still be the same in 10 years' time, or will we have found some means of registering graphemes in visual form and phonemes in the form of modulations? This already seems to be within the bounds of possibility, and the further question therefore arises: what part will mathematics then have to play? The answer will doubtless depend on two factors—on the success of mathematical linguistics, particularly of the current work on reducing at least some of the aspect of linguistic modes of expression to formal mathematical systems, and on the future importance of computation in the strict sense of the term (digital computation and logical computation) in the auto-programming and actual programmes of machine translation. Lastly, it is quite clear that the present-day algorismic systems whereby a French sentence can be converted into a succession of signals suitable for feeding into a machine offer obvious practical advantages and thus ensure mathematics a leading place in machine translation.

It is also clear that the linguist nowadays must become a mathematician and make use of the theory of sets as well as of graphs and Boolean logic, and even more, of course, of statistics, for language, because of its complexity as a social phenomenon, is not reducible to purely formal systems, but is amenable as such to statistical analysis.

To sum up, no general theory will help us for the time being to solve the

problems of translation. But every one of the scientific methods of analysis will be useful, and the research worker should be in a position to exploit them all. His special field, however, will be the detailed and accurate observation of a mass of facts with the aid of all the resources of mechanography and electronic computing.

His main attention will thus be focused on two objects—on the operation and principles of automatic digital computers, and on language-analysis methods designed to make it possible to 'feed' sentences into a computer with a view to obtaining equivalent sentences in another language. It is in this field, in fact, that studies and research already abound, if only because of the complexity of electronic computers, and because of the host of problems raised by their use for language work and the multiplicity of 'bilateral' language programmes (from English into Russian, say, or from Chinese into French). If we multiply the number of problems raised by a bilateral programme of machine translation by the number of possible bilateral programmes, the product is, in fact, enormous. The bibliography of articles on particular aspects of bilateral programmes is being endlessly enlarged, and provides a most fruitful and rewarding field for research.

As in the natural sciences, it is such special studies that count most, for the final and complete realization of machine translation depends essentially on listing, systematically studying and classifying a vast amount of linguistic data. The actual system of research may not yet have been perfected, but it can already be entirely visualized and its general structure defined independently of any general theory, just as the structure of chemical research was being defined in Lavoisier's day.

It follows that whatever progress is made towards machine translation, it will depend on the rational organization of work in two directions: the scientific training of staff for machine translation research, and the preparation of programmes for the machine.

The main purpose of the work done in these two directions is the creation of new social units, of a virtual symbiosis between man and a new tool capable of increasing his productive capacity manyfold.

But like any other creation of human society this tool cannot be improvised. A machine translation research team is not created overnight, even where a great number of machines are available for experimental work or for the actual production of translations.

THE FUTURE ORGANIZATION OF A MACHINE TRANSLATION RESEARCH TEAM

As indicated, we have urged the priority of the applied linguistics research workshop or laboratory over theoretical studies; for while the latter are of the utmost academic interest and are capable of illuminating our subject when required, the issues at stake are practical ones, and all research, not to

speak of all practical achievement, must be based on a study of language as a living reality.

How can machine translation research workshop be set up, and what will be the general principles governing team-work?

Wherever headway has already been made in this field, linguists and specialists in electronic digital computers have had to co-operate and therefore to learn to work together.

The first problem is to arrive at a mutual understanding and to 'speak the same language'. Three parties are involved—the linguist, the mathematician and the machine. Each uses a special language. The machine has its signals, its electronic impulses starting with input and resulting in output; language for it means an order, a precise and unambiguous instruction. The mathematician is trained in strict analysis, in the painstaking verification of his tentative assertions. The linguist is deeply conscious of the niceties and refinements of language and is perhaps better able to apprehend intuitively than to reduce his knowledge to precise formulae. If he is a translator, he often has had to give more thought to the irregularities and anomalies encountered in his work than to the vast body of repetitions and linguistic rules which he applies unconsciously. The linguist has made a practice of recognizing the exception, whereas the mathematician is trained to deduce the general rule from the individual cases.

The machine imposes a language, a choice, a standardization of the phenomena fed into it, and the mathematician or 'programmer' is its interpreter. The linguist has to learn to be selective, and to see a fact in the light of a single interpretation, and not of two; he must look for similarities rather than differences. Each has to familiarize himself with the other's methods, at least sufficiently to understand and respect them. The linguist contributes inventories of facts, while the mathematician tries to make them fit a system in which practically all phenomena can be reduced to formal rules. The former is the champion and interpreter of the recalcitrant stuff of language, the latter is the organizer who seeks to create rules from the diversity of facts.

The second problem is the choice of the machine to use. Should we wait for the ideal machine, the one best equipped to cope with complexities of language, or should we go ahead with what we already have?

The existing research teams have given us our answer. What Booth has done with the APEXC, and Panov, Belskaya and their colleagues have done with the BESM, Giuliano and Oettinger are doing at Harvard with the Univac I and Reifler, Micklesen, Wall and Hill at Seattle with an IBM-650, may be described as clearing the ground by demonstration and classification. Their work may be likened to pilot projects pointing the way to the machines and programmes of tomorrow. Why should we wait? It would mean condemning ourselves to waiting for ever, spurning the bronze axe and keeping to flint on the plea that the iron age will soon be here.

Besides, the team must be trained, and it will be more useful for it to use slow machines not too far in advance of the psychological reactions and knowledge of a group of persons who are acquiring their collective skills and forging their instruments. While linguists and mathematicians are thus training each other, the problems must be tackled one at a time, and there is no point in rushing ahead too fast or reaching too far.

The third problem is the choice of a source or input language and of a target or output language. Here, again, we need to have a clear view of the problem at the outset and avoid a dispersal of effort. The choice, for example, may be dictated by the more pressing need for translations from one given foreign language rather than another. It would probably be advisable for any team to confine itself initially to a single input language, and certainly to a single output language until it has been able to split up and form a second team whose members will already have been thoroughly trained in linguistic analysis for the purpose of machine-programming.

However, there is one need perhaps which has not been sufficiently stressed, owing to the fact that most of the work has been concentrated on translation into English. This is the need for synthesis, or the preparation of programmes enabling the machine to produce correct sentences in the output language. In the case of a language such as French, which is still relatively rich in stylistic turns of phrase, forms, and traditions, a degree of selection and simplification is necessary in order to compose the output language. From the means of expression available in the language, those have to be chosen which are essential, and those that are not discarded from the outset. The synthesis of a sentence by the machine means the reconstruction, from a sentence in a foreign language, of a French text that is intelligible and acceptable though it may lack polish. One of the very first tasks will be to determine just how rich or poor a glossary the machine will require for ordinary needs: for example, how many expressions should be stored in its memory in order to render the word 'no' in the sentence 'I have no bread' (e.g., *je n'ai* pas, point, aucun, miette, etc.). In many respects this process has much in common with the very interesting work being done on Basic French at the Ecole Normale Supérieure at Saint-Cloud.

The fourth problem is the choice of a set of texts or 'corpus' to be used for the analysis of the input language, and choice of methods for the analysis of this corpus. Every team, as soon as it is set up, will need to go into this problem and its choice will be guided by certain important considerations. Obviously, it will be preferable for the texts to be scientific, so as to limit as far as possible the initial extent of the vocabulary problems to be solved and to keep grammatical and structural problems within the narrowest bounds. A corpus based on one of the humanistic fields of learning or on general literature would raise too many problems and would set the team chasing too many hares. A series of texts on mathematics, astronomy, chem-

istry or any other exact science will be the most immediately useful and suitable for circumscribing discussions of minor aspects of linguistic work.

The fifth problem will be study of the work done by other teams. In the field of machine translation, no research team of any kind can operate at present without keeping in very close touch with what is being done by others —not in any competitive spirit but with a view to close collaboration and the mutual exchange of information. Bibliography and the scanning of journals and documents on all aspects of related work will be an absolute necessity and will cover a host of subjects such as the organs and working methods of computing machines, general and applied linguistics, and all branches of work relating to translation and language problems, together with the extensive documentation on bilateral programmes of linguistic analysis for machine translation purposes.

This list of tasks and problems, though far from exhaustive, enables certain conclusions to be drawn.

In the first place, the research must be on a collective and not on an individual footing. It involves a close acquaintance with different techniques and branches of knowledge which no single person can now possibly command as a basis for practical action. The only way it can be profitably organized is on a sufficiently large scale to make use of all the resources of a large body of specialists, most of whom would be working full time; although many marginal aspects of the over-all field of research could be entrusted to experts working as temporary advisers.

The second point is that this research is international, first because it is multilingual, and secondly because national frontiers are not coterminous with linguistic frontiers.

In the third place, the number of bilateral programmes is enormous. In theory, a multilingual country like the Soviet Union, which has 70 official languages would require $70 \times 69 = 4{,}830$ bilateral programmes. It will therefore be necessary to explore the possibilities of simplifying and extending the general coverage of programmes. In addition to immediate bilateral studies, this will involve a work of synthesis which, by bringing out the essential principles on which bilateral programming is based, will chart possible methods of preparing the blueprint for a prototype universal programme.

The fourth point is the emergence of a twofold task, outlining an all-round approach; there is need first for the rapid development of simple programmes (e.g., an English-French, or a Russian-French programme for scientific translations) so as to build up the stock of lexicographical and grammatical knowledge which those programmes yield as a valuable by-product; next, for the investigation of the theoretical bases of future programmes, with a view both to the formulation of multilateral programmes, and to the extension of programmes along more ambitious lines to cover such fields as humanistic studies, literature and poetry.

THE PRIMARY NEEDS: THE TRAINING OF RESEARCH WORKERS AND THE DEVELOPMENT OF EXCHANGES BETWEEN THEM

On the threshold of what may be the most marvellous scientific adventure ever conceived by man, one of the paramount requirements of any research —the need for research workers—is also clamant in the field of applied linguistics that now concerns us. The fact that we have to deal with disciplines which are wholly beyond the individual research worker's capacity and require the constitution of research teams in which each worker plays his part means that we must construct a vast block-diagram of research showing the place of each in the common task. One of the first areas charted is that for the training of the men who will then demonstrate their skill by rapidly proceeding to occupy the other squares in the diagram and finish the job.

A new programme of studies will need to be devised by an Institute of Applied Linguistics or some other appropriate institution operating at the higher education or post-graduate research level, and will involve the patient guidance of the members of this new profession towards practical achievements, the formulation of methods best suited to this type of work as distinct from traditional linguistic teaching, the combination of theoretical teaching and laboratory work, systematic corpus analysis and experimental studies in reducing sentences to their prime elements. The work of such universities as Georgetown, Seattle and Harvard, of the Massachusetts Institute of Technology and the Rand Corporation, and of Birkbeck College in London, the Steklov Institute of Mathematics in Moscow, the Institute of Precision Engineering of the USSR Academy of Sciences, and the University of Leningrad points the way forward and serves as a stimulus to linguists in other countries where university teaching follows more traditional lines, so that it is perhaps more difficult to merge the various disciplines. The need is for theoretical and practical studies, coupled with the constant and persevering checking of working hypotheses by laboratory experiments involving mechanical equipment and all the human and mechanical skills to serve it. The work of British, American and Russian research teams must be followed up and additions made to the constantly growing body of practical results which enlarge our theoretical knowledge thanks to the repeated testing they undergo on the machine.

If the training of research workers is necessity No. 1, the second is the organization of exchanges between them. The International Conference which was recently organized by Unesco was a confirmation of previous efforts in this direction rather than a beginning. International exchanges have been proceeding ever since 1952 when the first International Conference on Machine Translation was convened at the Massachusetts Institute of Technology with the help of the Rockefeller Foundation, and the scientists in the various countries are attentive readers of everything published on the work

of their colleagues. On all sides, a sense of fellowship in research is developing, and is clamouring for stronger expression—a sense of the need for closer and more regular exchanges which every new achievement intensifies.

Let us reflect for a moment on what is at stake. This is international research, in the real sense of the word, with the aim of facilitating and expediting scientific exchanges by surmounting the barriers that languages have erected between scientists and between men in every walk of life. The workers who are striving to penetrate the secrets of language with a view to machine translation are every one of them linguists, and specialists in the foreign languages on which they are working: their aim is to pave the way for exchanges, their aspiration is that nation shall speak peace unto nation.

This task, which, as we have seen, is a collective one on the national or regional scale, exceeding the capacity of single individuals, is essentially and inherently one for international team-work. The scope of requirements and of the research to be done makes it a fit subject for international co-operation, which alone will enable quicker results to be obtained through the division of labour.

How is this division to be achieved?

There are two well-defined phases in any general programme of translation from one language to another—analysis of the first language and synthesis of the second. We can already foresee how work will be divided up so far as synthesis is concerned. Each language study group (English, Russian, French, German, Spanish, Chinese and Japanese) can have only one synthesis programme at the outset since its aim is to convert algorisms into intelligible sentences in the national or regional language.

But the work of preparing this synthesis calls for the establishment of specific relationships between algorisms and linguistic means of expression: in other words, it involves analysis of the output language. This means that the German laboratory working on translation into German could compare and exchange its findings with French or English laboratories analysing German as an input language. The problems, while not wholly identical, are partly so.

This gives an idea of how exchanges covering a fairly wide field could be organized at various levels, and the same will apply to families of languages. The Romance languages will in all probability be able to use programmes with certain common features for translation from English or from Russian, and the same will apply to the Scandinavian or Slavonic groups.

Can such co-operation be aided or accelerated by organizing research internationally? This is a question we are bound to put even if it is impossible to give an authoritative or final answer.

An international organization responsible for promoting scientific and cultural exchanges could hardly fail to meet the wishes of scientists, if only by encouraging (even without organizing) co-operation between them. But

it is for the scientists to come forward in the first place and specify their needs and possibilities, with details of the assistance they desire in order to extend their work.

A start could be made with a clearing house for the exchange of information and programmes on a larger scale than is possible through the publication of material even in specialist reviews, and with a clearing house for the microfilming and exchange of those duplicated documents with which we are familiar but which are never circulated in an adequate number of copies. Another useful activity might be the organization of exchanges of students at the post-graduate level, study tours and visits by research workers and teachers to the various research centres. Lastly, in so far as work for machine translation impinges on neighbouring fields of research—and it does so to a very large degree, particularly in the fields of documentation, bibliography and abstracting—an international body, even on a small scale, should keep in touch with all those specializing in the mechanization of exchanges of documentation, scientific abstracts and bibliographies and keep its members abreast of any work in those fields likely to be of interest to them.

A minimum programme would be to publicize, in all appropriate cases, the work now being done, to communicate the findings of the most advanced teams to newly constituted ones, and to promote research by disseminating the results of previous work and saving newcomers from wasting their energies on research that has proved unproductive. It could be developed fairly rapidly in the form of technical assistance to countries needing to use up-to-date methods and international equipment for studies in applied linguistics for purposes of internal communication.

I am alluding here to the multilingual countries of Asia—to India with its 14 official languages and its sixty or so local languages; to Ceylon, where two or three languages are used side by side, and to the countries in which a European language of convenience exists alongside a vernacular language likely to develop as new sectors of society obtain access to science and education and need to cultivate and develop their language. International action in the world emerging from the second world war has taken different forms, of which technical assistance to the so-called underdeveloped countries is one of the most vital and practical. An international body of specialists in machine translation would soon have a part to play in this special field of international co-operation.

There is one last aspect of the foregoing problems on which a few concluding words should be said. International organizations have a tremendous need for translations, which are a heavy drain on their budgets—to say nothing at this early stage about the oral interpretation of speeches. The United Nations, Unesco, the International Labour Organisation, the World Health Organization and their regional bodies each spend an increasing

amount on the translation of speeches, documents and reports into a number of languages which are rather arbitrarily, and with difficulty, kept down to four or five. Hitherto, these organizations have paid little attention to the promises of the scientists who have heralded the coming automation of translation. Their attitude will change, for the promises are becoming more explicit. In conclusion, I must express my pleasure at seeing my translator colleagues and friends—international civil servants and members of the International Federation of Translators—taking an interest in the problems of automation and preparing themselves for the part they are to play in that respect.

CURRENT PROBLEMS AND RECENT NEWS

The scientific attitude: a possible misunderstanding

by

P. AUGER

P. Auger is Professor of Physics at the Sorbonne and was formerly Director of the Department of Natural Sciences at Unesco. He was Secretary-General of the International Conference on Radio-isotopes in Scientific Research, organized by Unesco and held in Paris from 9 to 20 September 1957, and of the International Conference on Information Processing in June 1959. He is also special consultant to the Secretary-General of the United Nations and the Director-General of Unesco.

To a great many people, man's entrance into the atomic era means merely the harnessing of nuclear energy. They picture the all-powerful atom as an entity far removed from everyday life, a kind of demon of the universe which scientists have managed to tame and put to work. This view is too limited by far and leads to a profoundly erroneous evaluation of the present trend in science. The present is indeed an atomic era but the term must be taken as applying to the whole atomic domain, that is, both the nuclei of the atoms, the atoms themselves and all the structures which atoms build by their combinations, more particularly molecules and crystals. The real feature of contemporary science is that the important events in it are those occurring at that microscopic level, whether nuclear, atomic or molecular, and that the phenomena observed, triggered and used by man on his own scale are simply the more or less direct repercussions of the first, toned down or amplified to a greater or lesser degree.

Further, modern scientists assent, implicitly at any rate, to the two following propositions: on the molecular scale everything is chemistry; on the atomic and nuclear scale everything is physics. At first sight these seem to be trite observations, even truisms. In reality, however, they are assertions with immense implications. The fact is that if, on the molecular scale, only the laws of chemistry (and of physical chemistry) come into the picture, this negates any interference on that scale by other types of laws or relations such as those which might be accepted, or suspected, by other sciences, e.g., biology or psychology. It is the same at the atomic level: since only the laws of physics are now operative, chemistry itself takes on a merely consequential status.

If we now combine the propositions in these two paragraphs, we find one of the basic attitudes of modern science. If phenomena on the human scale are ultimately repercussions on that larger scale of phenomena occurring at

the molecular and atomic level and if, in addition, the latter are determined exclusively by the laws of chemistry and physics, we are forced to the conclusion that the whole universe, including life, is governed by laws which are themselves no more than derivatives of these laws of physics and chemistry. In that case, the notion that laws of another type, other general principles, can influence the happenings of our daily life must be absolutely rejected. This firm attitude does not automatically require that the existing sciences, say, of psychology and biology, should be solidly tied now to physico-chemical phenomena exclusively, i.e., that the connexion asserted in principle should be clearly discernible everywhere. A vast amount of work still remains to be done and at present many discoveries made in these 'human level' sciences still remain disconnected, their roots in physical chemistry not having been fully established. Again, not all the laws of physics and chemistry themselves are yet known. However, in no case can new discoveries overthrow the propositions set out at the beginning of this paper.

This kind of attitude inevitably causes a profound shock to many sincere people. It appears needlessly dogmatic and, in addition, without sufficient foundation. If, it is often remarked, things were as the scientists claim, this should be much more obvious. We should no longer be confronted with those mysteries of life or thought extending far beyond the limits of our vision and offering us as yet no hope of plumbing them; therefore there can indeed there must, be something else as well. And that being so, why should not the action of these other laws, principles or forces originate at other levels, e.g., directly at human level, without our having to look for their genesis in the molecular or atomic sphere? It is perhaps exactly here that we find the real nub of the human problem of modern science, a problem which, in many respects, emerges as a misunderstanding.

There is, in fact, no question of the scientist's denying *a priori* the veracity of observations revealing actions of the kind instanced. All he is doing is to state his conviction that, if these actions are real, they could be traced back to a synthesis of actions at the microscopic level and would therefore conform to such universal principles as that of relativity or, at a less absolute but still high level of certainty, the principle of the conservation of energy, the principle of symmetry and the principle of entropy. One of the derivatives of these principles is the principle of the continuity, in time and space, of chains of causation, that is to say, the principle of the unbroken succession of phenomena passing on the energy and qualities of symmetry or dissymetry between the phenomena which together constitute the action under consideration, even if this manifests itself as a direct action at the human level. To the scientific mind no gap in time or space is tolerable, whereas to ordinary common sense this requirement may seem exaggerated in the extreme and with little justification. We know how greatly Newton was worried by the separated, unlinked actions which his discovery involved and

the enormous relief brought to physicists by Einstein's solution, with its restoration of continuity in space—and in time as well, since gravitation phenomena propagate at a definite speed, i.e., that of light.

Here it must be pointed out that the observed 'direct' actions on the human scale, adduced by the opponents of science, do not involve only physical objects and chemical substances, with man himself playing no direct part. The hazel wand or the radio-esthesist's pendulum are held in a man's hand; the tables at séances are under direct pressure from those present; the Rhine cards are read by men and are not photographed. This consideration alone practically rules out the question of the conservation of energy, since the energy balances can at all times be falsified, even sometimes unconsciously, by one of the individuals taking part in the experiment. You probably remember the story of the butcher who had lost his 1-kilogram weight and weighed the meat against his own fist, which he swore to housewives tipped the scale at exactly 1 kilogram. And when a suspicious customer brought along a 1-kilogram weight, the butcher had not the slightest difficulty in showing that what he had claimed was absolutely true.

A more delicate problem arises when we have to deal with entropy or symmetry, as it is no longer so easy to arrive at balances. Nevertheless, in this case too, what concerns the real scientist is to reach an exact appreciation of the value of all the transfers which may take place in the course of an experiment. I should like to give a few examples of the difficulty of the problem as it presents itself in the scientific study of objective phenomena, before going on to so-called paranormal or supranormal cases. Let us begin with fluid mechanics, the science which seeks to define the precise laws of the movement of water in rivers or streams of air in the atmosphere. These laws are reasonably well known but, in almost every case where it is sought to apply them to practical ends, they involve calculations so long and so complex that engineers and meteorologists usually prefer to resort, respectively, to direct experiments with models or to observations of such phenomena as nature presents. However, when a hydraulic engineer makes a scale model of the Seine estuary to study the effect of a particular breakwater on the bore, he is not for one moment calling in question that all that he is going to observe empirically arises from a summation of phenomena on the atomic scale, whose almost infinite number makes the total result indeterminable by our present means of calculation.

Incidentally, an additional difficulty presented by the laws of fluid mechanics is that situations very often develop which are of extreme sensitivity. Given a critical situation of this type, a minute modification of the mechanical conditions of the fluid mass under consideration may lead to profoundly different results. The same thing happens with random mechanisms like the Monte Carlo roulette wheels, where insignificant changes in the impulse given to the ball result in its landing in every number of the series

in an order impossible to foresee in advance; nevertheless no thinking perso doubts that the end result is the exclusive product of pure mechanics. Amon cases of this kind, some can be found which exemplify definite chains o amplification and my personal view is that these examples, in which th principle of operation is clear, can be used to interpret other cases in whic neither the process, nor even the principle underlying it, are clearly ap parent. Here I am thinking specially of the biological phenomena whic occur in highly complex cellular systems such as the brains of higher animals.

Indeed, it is from biology that I propose to take my second example and specifically from the domain of genetics and embryology. A scientists ob serving the development of living beings from their reproductive element naturally seeks to determine in detail the mechanism whereby a minute germ within which only very few details can be distinguished by our instruments nevertheless produces with complete reliability a fish, an elephant, or a fly as the case may be.

The important point for us is that the biologist does not question th existence of such a mechanism, a step-by-step sequence linking the charac teristics of the developed entity to 'something' which is present in the germ Today there is a very widespread belief that this 'something' is in fact a rigid arrangement of molecules, embodying the quintessence of a specific complex structure from which this structure itself can be obtained by th application of automatic mechanisms. Looked at from this point of view genetics is rooted in a specialized chemistry, and the chains of events con necting that chemistry to general biology on our own scale admit no di continuity either in time or space. Any interval during which (in the cas of time) or within which (in the case of space) this structure is not physicall present automatically destroys any chance of the chain's running on and thu brings the development to a full stop.

To this picture of the mechanisms connecting the molecular and th macroscopic worlds a further essential detail still needs to be added. Thi is the direction in which the link-up occurs, for it is a one-way proces working exclusively from the smaller to the greater. In the example jus given, from genetics, we saw that the microscopic order governed absolutel the macroscopic. This link-up conforms to the laws of Mendelian heredit with all that the Morgan school of biologists has been able to add to them A link-up occurring in the opposite direction would be of the kind postulate in earlier days by Lamarck and would involve the possibility of specific reactions by the macroscopic forms of fully-developed entities on the micro scopic order of their germinative elements. It has never been possible t adduce certain proof of the existence of this type of connexion and it migh well be that it is contrary to some natural principle. This is not to say tha evolution is not subjected to the influence of the macroscopic living condi tions of the developed entities but it does mean that the influence is exerte

only through the indirect mechanism of selection. There is thus no direct specific action by events at the large-scale level on the small-scale order.

Other examples could be found in the domains of medicine, of chemistry or of psychology. Throughout antiquity and the Middle Ages men believed in a system of medicine working on contraries or analogies with remedies selected by comparison of their effects with the general symptoms observed. It was only when the microscopic causes of disease were identified that the first effective system of medicine was established by Pasteur, founded on the microscopic specifics represented by antibodies. Of later years the manufacture of these specifics is no longer left to the sick organism itself, acting as a microscopic factory: it has been possible to pass the task to other living beings such as the moulds (penicillin). Finally, in the case of the sulphonated antibiotics, it has even been achieved in the laboratory through a detailed synthesis of molecules exhibiting structural arrangements sufficiently like those of some of the molecules forming bacteria. The use of general antiseptics can in no circumstances be comparable in effectiveness: as their specificity does not extend down to the microscopic level, they act indiscriminately on a whole enormous category of organic molecules and kill the corresponding cells in the body of the sick person as well as those in the invading microbes. In chemistry it is enough to instance catalysis to see the same microscopic specificity at work in the processes of molecular synthesis.

In psychology, if we agree to ascribe the emotions and passions to the macroscopic domain and concepts and ideas to the microscopic, it becomes fairly clear that the first group are under the control of the second but not vice versa. Obviously, we must first have cleared out of the way the myths of the poet in the frenzy of creation, or the heat of battle inspiring a general to invent a stratagem. The passions often do provide the drive and probably may sometimes speed up thought, but the thrust is in a general direction and the detached creative process involving the making of combinations of ideas, with choice between, and exact arrangement of them, requires detailed mental activity which sometimes takes a considerable time and may in part be unconscious. Briefly, the modern concepts of scientific psychology cannot but condemn the theory of inspiration (which is no more than a disguised form of the theory of revelation), which postulates the instantaneous emergence in perfect form of a new structured whole resulting from the impact of some unstructured external impulsion—whether it be faith, nostalgia or love.

These specificities and this ordered succession represent aspects of 'information', a new physical parameter comprehending in itself a mass of concepts and ideas scattered over the sciences, the technologies and the arts. It represents the essence of every idiosyncratic structure, of every arrangement of parts which might be otherwise and which accordingly represents a choice between various conceivable combinations. If I put down 100 francs

to play pitch and toss, the result announced after each throw is the unit of 'information', since we are dealing with a single alternative. If I have a ticket in the National Lottery, the 'information' given when the draw is made equals the number of tickets sold, since I might have bought any one of the others instead of the one I did buy. Comparing the chances of winning in these two gambles, it can be seen that 'information' can be related directly to very definite values.

Turning to biology, we might say that the 'information' contained in the chromosomes of the germ of an animal is called on during growth to determine the characters of the adult. We know that the smallest change in the fine structure of these chromosomes has repercussions on the conformation and functioning of the fully developed being. Biologists already have some ideas about how this 'information' is physically incorporated in the chromosomes: it is probably something very close to the lottery tickets, i.e., a very long sequence of figures.

We can then understand, or at least glimpse, the 'why' of the unidirectional process whose existence has been pointed out in the relations between the microscopic and the macroscopic. If the 'information' in the chromosomes is modified by an action on the molecular scale, the large-scale result is altered, just as the alteration of a figure in your ticket may make the difference between your winning or losing. But no large-scale reaction such as rage or satisfaction can make any change in the figures of the draw or the figures on your ticket—just as a 'gross' action on the developed entity can produce no effect on its germinative 'information': at most you can kill the individual, which is the same as tearing up your ticket. The fine-scale action alone can produce a particular fine-scale result and thereafter set off the extended consequences issuing from it.

It should be remarked, incidentally, that, in the large-scale consequences, the initial information may subsist in all its particulars. This is not the case with the lottery, but it is with biological growth, and nevertheless, even there, we cannot put the machine into reverse and cause a change in the macroscopic 'end-structure' to react on the microscopic structure which determined its growth. Each time experiments or observations have appeared to imply such a reverse action, it has been possible to show that, in fact, this was an illusion or was due to the action of an hereditary or selective mechanism which had not been recognized initially. Thus in the extremely interesting research lately carried out first on bacteria and then on ducks, in the course of which the hereditary characters of the species seem to be permanently modified by the action of certain nucleic acids, there is no doubt that the action in question starts on the microscopic scale by the insertion of a new element into the structure of the nucleus alike of the cells of the body and of those constituting the germs of the next generation. This gives us two series of independent consequences, one affecting the characters of

the adult and the other the hereditary traits carried by its germinative elements, so that the two kinds of consequences are not necessarily present together and in strict parallel.

Consideration of the new physical parameter of 'information' will make it possible to exhibit the misunderstanding mentioned at the beginning of this paper in a specially acute form. Let us take, for instance, the simple necessity for temporal and spatial contiguity between all the links in a chain of phenomena proceeding from a given structure to another which is their consequence. If we take the case of a message conveying information, we must conclude that the signals of which the message consists do in fact travel from the transmitter to the receiver without a break and carry in themselves —physically and at every moment of their passage—the information in question in their very structure. This conclusion agrees completely with all scientific observation and communication techniques. Now let us look from this angle at the experiments made over the last few years in the United States of America in an attempt to demonstrate the existence of 'extrasensory' perception in certain human individuals. The statistics cover millions of cases; the calculations of probable errors appear to have been worked out with the greatest care; and yet the vast majority of scientists reject the conclusions reached, almost—it might be said—at first sight. Why? First, and basically, because those conclusions require a departure from the principles of continuity in the transit of the information. Let us see how.

In the first experiments, an operator turned up the cards in a pack out of sight of a subject who 'guessed' them as they were turned. However improbable it may seem to the physicist, it could be granted that messages of an unknown nature were passing between the individuals. Then in later experiments, the subject was put farther away in a neighbouring room and ultimately even in a different town. The passage of the messages thus became progressively more and more improbable but not absolutely inconceivable. At this point, however, the subjects started to 'spot' the cards either some time after, or some time before they were turned up. Seemingly the theory of information had been conclusively breached . . . but there might be a way out, supposing the subject to have long-distance cognition of the order of the cards in the pack before they are scanned by the operator.

Let us analyse this more closely. Messages must pass between the cards and the subject. These messages come from the images on the cards. Let us grant this, however extravagant it may seem. But how can these messages give an indication of the order in which the cards lie? The pack fits closely together; the images are practically one on top of the other; the lie of the pack in relation to the subject is not defined and the radiation of the messages must go on uninterruptedly from all the cards at once, since the individual cards cannot be 'cued' that it is their turn to be considered. Hence the subject has no means of naming them in their order. When we come down to it, it

would be no more surprising if the cards had been named even before they were shuffled and the pack put together!

Many sincere people have been surprised at the little interest shown by the scientific world in these experiments and the lack of eagerness on the part of scientists to carry out checks, easy though this would be.[1] Indeed, a new kind of message would be nothing to frighten them but they would still demand that temporal and spatial continuity be maintained in the passage of the information. If it is not, the scientist considers that the chances of errors or trickery are too great for him to waste his time showing them up. These remarks of course apply equally to all alleged cases of premonition in breach of temporal continuity, and to the pendulum searches of the radiesthesist on maps or ground plans, which disregard continuity in space. When we read the works of the serious and patient writers who have described their experiments in extra-sensory perception, it is at once evident that they have concerned themselves but little with tracing the required continuity and contiguity in time and space over a hypothetical chain of causation. They are apparently ready to assume that the mode of action at a distance which they are studying is not bound by these principles: after all—they say—there is still so much that we do not know and, anyway, there is always the possibility of the principles of physics being exploded by some new discovery. What they do not realize, however, is that here we are up against something much more fundamental, so that, before we can shake a basis as solid and essential as that provided by the great principles of conservation, the principle of relativity, the principle of symmetry or the principle of uncertainty, we need more than a few coincidences in drawing cards.

And after all, between ourselves, do you not think that modern developments bring us enough marvels, established and usable marvels, for us to be able to assuage the old longing which subsists in us from the days when we listened to fairy-tales, without the aid of weeping or bleeding statues, furniture moving by itself, pendulums or wands. There is no room in our universe for the old impedimenta of a time now past, a time when we did not yet know the royal road which is now leading us forward, unfaltering, from the electron to the galaxy.

1. Thus no one has ever tried to have the subject 'perceive' whether cards were blank or bore usual designs!

International Conference on Information Processing

Paris 15 to 20 June 1959

by

S. DE PICCIOTTO

S. de Picciotto, is an expert on scientific questions relating to the use of large electronic computer systems. He was appointed by the Department of Natural Sciences of Unesco, as editor of the Proceedings of the International Conference on Information Processing. He is now an engineer with the IBM Corporation (France),

INTRODUCTION

Since the earliest times man's struggle with nature has had two coexistent and sometimes complementary aspects. The first is an opposition of forces: the physical strength of man versus the power of natural phenomena. The second may perhaps be described as 'organization', and is partly bound up with the social nature of the human animal: it is through their greater or lesser cohesion—their mode of organization—that human communities have learned to survive since time immemorial. This traditional duality recurs in modern science and technology.

On the one hand, there is the fact of the continuous development of research and the ever more spectacular achievements aimed at discovering and exploiting more powerful sources of energy and more effective instruments of production: man's increasing material requirements have given birth to mass production, which science daily provides new ways of effecting, automation being the most significant of the advances made in that respect.

On the other hand, there is the fact of the development of social humanistic and political sciences resulting from new methods of economic organization, scientific management and actual planning. Mathematical methods are being devised for research and rationalization. The use of statistics is making it possible to deduce general laws, while methods collectively known as 'operational research' are being applied with a view to working out optimum values for economic management and policy-making.

These organizational methods, like pure and applied scientific research make use of quantitative concepts expressed in figures and symbols, the processing of which becomes more complex every day.

Statistical data on economy and management, as well as equations in nuclear physics, can be put to use only after a long process of elaboration, calculation and interpretation of results; and the growing volume of such

data, and the growing abstruseness of the theories which make it possible to use them accurately, render the information material derived from measurement or observation practically inexploitable by normal methods.

The automation of industrial production processes has to be matched by the automation of information processing if scientific progress and the rationalization of economic organization is to continue.

Information processing is thus a matter of prime importance for the future of mankind. The question is no longer one of more or less efficient processes for accounting or for technical calculation but of devising methods in keeping with the atomic and interplanetary civilization of tomorrow. The issue is one of universal moment.

Many countries have already made great progress in research work and results in this field; there was thus an obvious need to get together and take stock of what had so far been achieved. It was with this consideration in mind that, in 1957, the American Joint Computer Committee asked Unesco to organize an international conference on the subject.

At its fiftieth session, the Executive Board of Unesco approved the principle of convening the proposed gathering, and this was confirmed by the General Conference of Unesco at its tenth session.

Its decision set the Secretariat a host of tasks, the most important of which understandably related to the actual content of the material contributed. The difficult work of delimiting the subjects to be covered, followed by the selection of the papers to be read from among those submitted by international experts, was done by a group of consultants consisting of the following: S. N. Alexander (USA), I. L. Auerbach (USA), J. Carteron (France), J. Coales (United Kingdom), S. Comet (Sweden), P. Dreyfus (France), E. Durand (France), A. Ghizzetti (Italy), M. Goto (Japan), A. S. Householder (USA), C. Manneback (Belgium), P. Namian (France), D. Y. Panov (USSR), W. L. van der Poel (Netherlands), R. de Possel (France), R. Rind (France), C. S. Scholten (Netherlands), K. Steinbuch (Federal Republic of Germany), A. Walther (Federal Republic of Germany), A. van Wijngaarden (Netherlands), M. V. Wilkes (United Kingdom), and H. Yamashita (Japan).

Out of about 170 papers submitted, 60, exclusively theoretical in nature, were selected.[1] Adopting the view that the practical aspects of information processing should be subordinated to the more fundamental problems of theoretical research to be discussed at the conference, the consultants decided to eliminate certain papers of very great interest.

The 60 selected papers were printed in advance and abstracts of them translated into five languages: English, French, German, Spanish and Russian. The conference programme covered three types of activity: (a) plenary meetings, where the 60 selected papers were discussed; (b) symposia, where

1. A list of these papers appears at the end of this article. References to them in the article are indicated by the number given in brackets.

data were exchanged concerning current research in various specialized sectors of information processing; and (c) evening lectures open to the general public.

Side by side with these proceedings, the National Commission of the French Republic for Education, Science and Culture organized a series of very successful meetings, in collaboration with the Unesco Department of Social Sciences, on the subject of 'the use of mathematics in the social sciences'.

The main subjects discussed by the conference were the following: methods of digital computing; logical design of digital computers; common symbolic language for digital computers; automatic translation of languages; pattern recognition and machine learning; collection, storage and retrieval of information; computer techniques of the future.

The conference, which was non-governmental in character, was attended by 1,772 participants (in their individual capacities) from 37 countries, and was presided over by Professor Howard H. Aiken (USA), creator of the first large-scale general purpose machines Mark I and Mark II.

An illustrative exhibition of specialized equipment was organized in connexion with the conference with the help of the Association for the Development of Information Processing, under the direction of Mr. Max Hermieu, and was presented from 13 to 23 June 1959, with 28 exhibitors from eight countries displaying equipment incorporating the latest techniques.

Our aim here will be to describe the essence of the work of this first conference to be held on the subject, for it is not our intention in these pages to involve the reader in the details of papers of so highly technical a nature. As to the experts, we must crave their indulgence for the sketchiness of this report, which is solely designed to give some idea of the present status of a first-rate scientific discipline.

METHODS OF DIGITAL COMPUTING

Numerical analysis is a difficult science in which specialists come up against particular snags not found in the other mathematical disciplines. The latter usually express their results in written formulations in which the number-symbols have a strictly absolute significance. The mathematician notes that the area of a circle of radius r is πr^2. The numerical analyst goes further and specifies that 'if the value of π used in the calculations contains a relative error $\frac{\Delta \pi}{\pi}$ and the measurement of the radius is made with an error $\frac{\Delta r}{r}$, the result of the calculation of the area will be distorted by a relative error

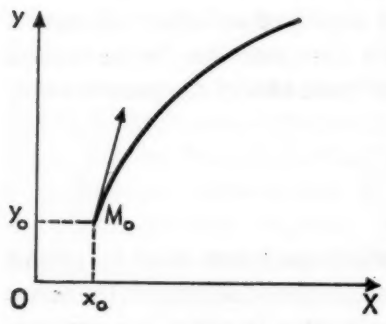
less than $\frac{\Delta\pi}{\pi} + 2\frac{\Delta r}{r}$'. Turning next to the question of figures, he continues: 'Where $\frac{\Delta\pi}{\pi} = \frac{1}{10{,}000}$ and $r = 1$ metre to the nearest millimetre, the area S to be found is between 3.125 sq.m. and 3.1487 sq.m.'.

This descent from mathematical formalism to physical measurements (which are readily assessable in this elementary example) is far from being always controllable.

The work of the conference showed that research workers are at present preoccupied by four main themes.

Differential equations

The simplest examples that can be given of these are in the field of mechanics. Let us consider a material point in motion which is subjected to certain known forces: mechanics teaches us that if we known the position and speed of that point at a given 'initial' instant, its later movement is also known. If we now consider the movement of a point on a plane, it follows from the above that it is sufficient to known a point on its trajectory and its tangent at that point in order to be able to determine it completely.



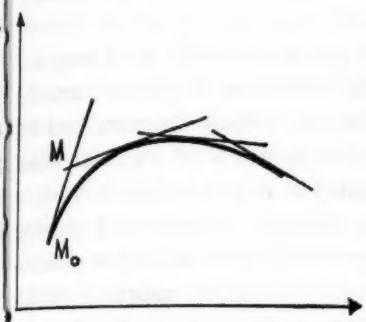
More specifically, the movement of the point is fully known where the following are given: (a) the permanent relationship between the co-ordinates x and y of the moving point and their 'derivatives' in respect of time (this relationship being known as a differential equation); (b) the values of these co-ordinates and their derivatives at the initial point.

Where a formal process is known (the differential equation being given) whereby x and y can be deduced from the latter at any time, the problem is practically solved; but the substitution of numbers for letters gives rise to the problems of errors already considered.

Where no formal method of solution is known, use is made of approximation methods, the validity and applicability conditions of which are the subject of highly complex studies.

The most common example is that where (the tangent at the initial point M_0 being known) a point M is taken on that tangent near enough to be regarded as falling within the trajectory. It is then possible to determine the

tangent at that new point by calculation. It is possible in this way (with a certain degree of error) to find all points on the trajectory by closer and closer approximation.



It is obvious in this case that the accumulation of error at every step may produce a line very different from the trajectory sought. Nor does reducing the steps, though permitting closer approximation at each point, solve the problem, for the resultant increase in the number of points also increases the number of calculations, itself a source of accumulation of error.

These few considerations give an idea of how difficult these questions are. The point is, however, that the solution of differential equations is of importance for a wide range of scientific, technical and economic problems.

Some of the papers presented at the conference related to 'initial condition' problems. It would seem that in the case of 'separate-step' methods the specialists [1] are now advocating the reduction of differential systems to the 'canonic form', when higher order derivatives are present.

As far as the mechanism of the propagation of error is concerned it is possible by theoretical means to define its limits. But these evaluations are usually exceeded in practice. The studies on this point are both theoretical and experimental [2].

What has been said regarding differential equations indicates that it is usually a question of determining a curve, whereas the simple cases referred to involved the determination of a line in a plane or in normal three-dimensional space. While still keeping to these, we can single out another type of calculation with which numerical analysts are at present concerned: partial differential equations.

Partial differential equations

These equations relate to problems where—in contradistinction to those mentioned above—the determination relates to 'surfaces'. Cases in point are the study of the deformations of an elastic membrane stretched over a rigid frame representating its 'boundary conditions' and subjected to certain mechanical stresses, or the study of liquid flows and waves.

Each of the points of the physical medium concerned in these cases depends on two or more parameters, and the function to be found in respect of these variables enters into the calculations, as do also its 'derivatives' in respect of each of the parameters concerned.

In the simple case of two-dimensional problems, it is no longer a question of determining points on a trajectory, as in the problem referred to above, but of attaching to each point of a given plane field a numerical value resulting from the calculations. This problem is similar to that of determining contour lines on a map.

Partial differential equations raise difficult problems which are a long way from being solved. These equations are very important in physics; and the propagation of heat or of electromagnetic waves, atomic structure and the laws of hydrodynamics and aerodynamics are only a few of the fields which directly depend on them. Research on the subject is proceeding in a whole series of countries, and the ever-increasing 'storage' capacity and growing speed of operation of large electronic computers is now making it possible to face up to the difficulties to which the multi-dimensional nature of partial differential problems gives rise. But the obstacles are not confined to the technical aspects, and much remains to be done on the theoretical side.

The methods of approach are various, and obviously depend on the type of equations concerned. In some cases the latter can be reduced to ordinary differential ones [3], but it is often convenient to substitute discrete variations for the infinitesimal variations implied in the concept of 'derivative'. Partial differential equations thus become 'finite difference equations' [4, 5]. The problems are thus reduced in particular cases to the solution of linear systems or become amenable to certain iterative methods [3, 4, 6].

This, in turn, raises important problems of 'convergence'. It is necessary to make sure first of all that the iterative process concerned leads in fact to the desired solution, and then that the errors accumulated during the calculations are not propagated in such a way as to be prohibitive [3, 4].

The evaluation of possible error in calculations has been the subject of theoretical studies [5] which bring in the latest concepts of abstract mathematics. These, too, reveal the preoccupation common to all sectors of numerical analysis. The difficulty arises from the fact that in this field every case is a special one, and that only rarely can general conclusions be drawn. Algebraic and matrix processes are the subject of searching studies [7, 8, 9], in which the statistical and probability approaches seem to hold out great promise.

Algebraic and matrix processes

The first degree equation systems, familiar to secondary-school children, present formidable problems from the standpoint of accuracy where the number of equations increases; and for certain so-called 'ill-conditioned' systems a tiny error in a single coefficient so affects the solution as to preclude its significant appraisal. What makes these linear equations important is the diversity of the problems of which they are the expression. The sets of

coefficients of these systems or 'matrices' play a decisive part in the calculations of molecular and atomic structures, as well as in statistics or economics where problems of correlations are concerned.

Certain problems relating to 'eigenvalues' also remain unresolved, and research on that subject was discussed at the conference [9].

The same applies to the evaluation of elementary and special functions (sine, cosine, logarithms, Bessel functions, etc.), the accurate and more rapid determination of which is a matter of ever-growing importance, and the conference discussed the possibility of rational approximations by means of polynomial and continued fractions [10].

In the final analysis, all these methods and abstract calculations relate to specific and concrete physical problems. For example, one paper on diophantine algebra [11] concerns a queue problem, while another on multidimensional spaces [12] is directly connected with questions of economic management.

Operational research

The importance of 'operational research' for the definition of optimum values in economic, administrative, strategic and political fields no longer needs stressing. The traditional approach had resulted in a rule of thumb whereby the desired optimum could be found by a so-called 'simplex' method on the basis of a set of conditions expressed in the form of a group of linear equalities and inequalities. The following example will give the reader some idea of this method: proceeding from the polyhedron whose surfaces are defined by the above-mentioned linear equations, the aim is to find the summit closest to the plane corresponding to the required economic condition by proceeding along the edges of the polyhedron.

'Proceeding along the edges' is a painfully protracted operation; and the tendency of recent research has been to substitute an optimum value within the polyhedron for one located on the contour of the field.

Generally speaking, numerical analysis gives the appearance at present of being a discipline in full development but not as yet clearly displaying any dominant features. It is not easy to describe it in comprehensive terms: as we have already stressed, each problem has its special nature, and the essential characteristic of this discipline at the moment is its increasing tendency to become an experimental branch of mathematics through the use of electronic computers.

The coupling of the words 'experimental' and 'mathematics' may shock certain people, but surely the time has passed when anything should be able to surprise us.

LOGICAL DESIGN OF INFORMATION-PROCESSING MACHINES

Before reporting on the work of the conference in this field, it is as well to recall certain basic points concerning the organizational structure of conventional data-processing systems.

The information, for all problems, comes in the form of basic documents which are then subjected to a series of complex processes: sorting, merging, calculation and classification. The most striking features in this connexion are the volume of information, the number of arithmetical operations and the intricacy of the situations arising at any given moment and requiring a large number and variety of corresponding logical decisions.

Hence a machine designed to handle such problems will have to include a large-capacity memory unit for storing the data to be processed, and a high-speed computing unit capable of carrying out arithmetical operations and also of performing logical decisions.

In addition, it should include a control unit—a group of organs designed to ensure the auto-determination of operations and the interpretation of the user's will expressed in terms of the machine's own language.

It is the existence of this internal language which is the basic characteristic of electronic information-processing machines.

All calculations, irrespective of the means whereby they have to be processed, involve practical preparation, with an execution chart and plan representing the various stages of the work in view—a programme, in other words.

In the case of a machine, the programme takes the form of a set of instructions expressed through the medium of its own language and recorded in its memory, whence the control unit will have to retrieve them for interpretation purposes.

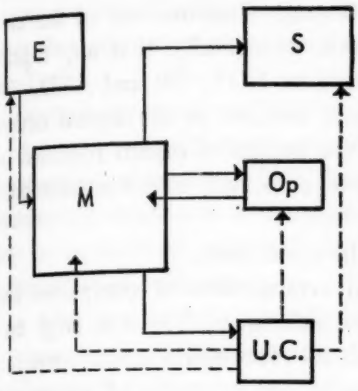
It is thus possible to envisage two main categories of information for processing by computer: the factors (data) involved in the various operations, and the instructions determining the course of those operations.

Lastly, another unit has to be added to these various organs to enable the information to be fed into the machine, and another for supplying the processed results to the user.

Such, then, is the general lay-out of computers; but there are a number of questions that arise.

A machine of this kind is designed mainly for digital processing. Both economic and logical grounds argue in favour of trying to reduce the number of physical elements in the computer design to the minimum, with figures represented by different arrangements of 'bits' or electrical impulses, corresponding to the sequence of changes in the system in accordance with the machine operations.

The calculating organs are of the electronic type, functioning between



General lay-out of a conventional computer

E input unit for feeding information into the machine
S output unit which delivers results
M large capacity memory unit for storing data and instructions
Op arithmetical and logical operator for processing the data of the problem to be solved
UC control unit—analyses the instructions received and, depending on their nature, either sets arithmetical or logical operations in train in Op, or sets E or S in action.

two mutually exclusive stable states. This method of functioning permits 'all or nothing' effects, thus providing the basis for a binary notation system which is extremely advantageous because of its directness: 1 indicates the presence and 0 the absence of an impulse, and all numbers can be represented by combinations of 1 and 0.

The main circuits for processing these impulses are logically simple. Their conception and calculation belong to a special form of algebra known as Boolean Algebra, after the English mathematician Boole, and they relate essentially to the following operations: logical addition, known as the OR operation; logical multiplication, known as the AND operation; negation, where x corresponds to its opposite, represented by the symbol x or $(1\text{-}x)$.

y \ x	0	1
0	0	1
1	1	1

OR operation. The two Boolean variables x and y (which may assume the values 1 and 0) are made to correspond to the Boolean variable z, in accordance with the diagram shown opposite. In other words, it is necessary and sufficient, if z is to equal 1, for either x or y to equal 1 also.

y \ x	0	1
0	0	0
1	0	1

AND operation. z can equal 1 only if both $x = 1$ and $y = 1$ (see diagram).

x	0	1
$\bar{x}$	1	0

NOT operation ('negation' or 'complement'). The correspondence table between x and x is immediate.

All the complete circuits are obtainable through combinations of the elementary circuits described. It can be shown, incidentally, that any logical function can be expressed by the three operations NOT, OR and AND.

The above points underline two important features of the logical design of computers: the general arrangement of the groups of organs regarded as functional units (inputs, storage, programmes, etc.), and 'micro-organization' at the basic circuit level.

All these questions were discussed by the conference.

The development of the general logical arrangement of computers has been closely related to that of construction techniques. The first large machines, such as Mark I, for example, with its electromechanical operation and its comparatively limited storage capacity, are incapable of storing the complete sequence of instructions related to a given problem: these machines are of the 'external programme' type. The programme, coded in the form of punches in a paper tape or on cards of the type used in accounting machines, is gradually 'absorbed' into the machine, which consecutively executes the orders received. The result is to reduce the speed of operation through the slow action of both the computing and the input units.

The speed gained by the computer circuits consequent on the introduction of electronics would have been all to no purpose but for the invention of new methods permitting the rapid input of instructions; and the way this problem was solved was by increasing storage capacity and enabling instructions to be pre-recorded in a medium allowing of immediate consultation. The concept of the 'stored programme' was thus born.

This historic example helps us to understand present trends, with ever more spectacular performances by input-output devices designed to permit, among other things, the continual use of the machine at its maximum computing speed.

Until quite recently, the conventional view was that computing could not take place during an input-output process. This led to the creation of 'peripheral' units, with the preparation, independently of the main computer, of magnetic tapes from cards or paper tapes, or, conversely, the printing of results and punching of cards from magnetic tapes.

The present approach (and this will apply even more forcibly in the future) is in favour of the organization of simultaneous-action multi-purpose machines, working on the principle of time-sharing, but also on that of space-sharing, with an apportionment of the 'space' available in the memory and a clearly defined allocation of particular groups of memory cells to various sections of the work.

Both theoretical and practical studies designed to solve these distribution problems were presented and discussed at the conference.

Suggestions were made for employing a monitor routine for controlling priorities in the use of the machine's various organs [13]; while coded tech-

niques for interrupting the operation of individual organs in the computer were also advocated as a more specific answer to the 'input-output' problem [14].

Another approach was in favour of constructing machines comprising specialized units with independent programmes, one of the units to handle input-output operations [15]. The advocates of this idea, who have rather humorously applied to it the term 'sympathetic programming', hope in this way to co-ordinate different elements which would remain more or less independent of each other.

A different idea is to include, in the circuits for decoding instructions an 'urgency' discriminator reacting in accordance with pre-assigned priority coefficients [16].

Some specialists have boldly adopted a more radical stand and advocate building machines which really consist of several computers sharing each problem among them [17]. A primary computer—a high-speed general-purpose unit—carries out most of the calculations; while another machine handles special marginal but necessary operations, and a third is the link with the outside. The information is transferred through a store common to all three computers.

Other constructions have still further analysed the detailed shape of the information handled by machines in general and have worked out the idea [18] that the data processed in a general-purpose computer is of four quite distinct types:

1. A form of notation proper to scientific calculation, where accuracy and speed in computing are the major requirements. The processing entity must in that case be a number comprising the maximum number of significant figures, with automatic specification of the position of the binary or decimal point. This is the 'floating point' notation.
2. A representation of numbers, suitable for administrative and accounting computation where the processing entity—the 'word'—is essentially of variable length, in which the order of magnitude is generally well-defined. The process is that of a 'fixed-point' calculation.
3. A type of fixed-point information where the 'word' length is also fixed and which is generally employed for processing instructions even when they intervene in the programmes.
4. Lastly, there are the varied forms which information may take when it is set up for the machine to put out.

This analysis culminates in the idea of a machine equipped with a number of specialized computing centres each processing a particular form of information.

These new ideas and concepts have now reached varying stages of maturity and it is not easy to forecast future developments in this field; opinions are hazardous, in fact, for reality may soon outstrip the boldest conjectures.

Micro-organization

In studying information processing at the level of elementary operational units, the result of the new techniques has been to compel a reconsideration of principles. The present methods of approach are varied.

The studies just described [18] have led not only to a new conception of general structure, but also to that of the simultaneous use of various modes of information flow.

Fixed-length 'words' circulate 'in parallel', or, to put it in more readily appreciable terms, each component digit in a number is transmitted from one of the machine's organs to another via an individual conductor. Hence the time needed to transfer the complete word is as short as that needed to transfer each digit.

Variable-length words travel 'serially', the same channel being used consecutively by all the digits in the word transmitted. Each type of information has its own and perfectly adapted mode of circulation.

The user thus has immediate access to the most elementary operational entities, including even isolated binary digits.

Progress in improving the micro-organization of computers also depends on the techniques used in arithmetical operations.

By overlapping in time [19] the operations of addition and shift in multiplication, by applying a special method of handling the remainder in a division operation, and by instituting microprogramme circuits, electronic computations can be speeded up considerably.

Another technique for achieving that end is based on a study of carry propagation during computing operations [20].

Special logical circuits are continually being added to the already wide range of those used by constructors.

Address selection switching in order to direct data to a particular memory cell [21] has also been closely studied and has made it possible to use an error-correcting code with a minimum of hardware.

One interesting generalization of elementary logical circuits has been the introduction of 'elements on majority decision principle', defined as organs for which a 'majority of binary values of their inputs decides a binary value of their output' [22].

A paper on base-three digital circuits, challenging the very principle of the numeration system to be used, describes the classic operators AND, OR, NOT, etc., in a three-valued system [23].

All the above papers dealing with the elementary circuits of computers are obviously determined down to the smallest details by the discontinuous nature of the numerical information handled. Mathematical observations [24] on this fact lead to a system of representation of great practical interest.

This 'discontinuity' has also led to the conception of systems for

converting continuous physical measurements (rotations of movable shafts in a machine, for instance) into figures which computers can use, and a new coding system on these lines was presented at the conference [25].

The essential feature of all this work is that it is based on the desire to construct even larger capacity machines.

Computers do, however, present a financial problem which leads constructors to try to produce average or small capacity machines offering a maximum of technical advantages.

One way of achieving this end is to simplify the machine's circuits and functions to the hilt so that as many operations as possible, even elementary ones, are carried out by means of microprogrammes [26].

The financial aspect can be considered from another angle. Where the budget is limited [27], the machine must be tailored to suit that budget. There is no reason why interesting results should not be obtainable despite limited funds.

The conference thus afforded specialists an opportunity of comparing their ideas and achievements at all levels of computer design; and notwithstanding the general unity of certain principles, the striking thing was the diversity and wealth of the ideas advanced: every one of these machines, differing in origin and capacity, had its own systems of codification and operation.

The question arises of whether the user, faced with so many computers will not run the risk of getting lost in the 'babel' of language peculiar to each machine. For each of these wonderful creations has its own internal language, consisting of the instructions for the operations which it is to carry out and which are all controlled by means of these magic 'words'. Would it not be possible to standardize this vocabulary and convert it into a language common to all the machines alike? [28]

A COMMON LANGUAGE FOR ALL MACHINES

The instructions, in a machine, are numbers presented in the form of binary impulses. Codes and addresses numerically written in basic notation thus fixed by construction are called absolute. The machine processes and interprets instructions codified as absolutes and ignores all other codification forms. Thus the transfer of mathematical formulae representing a problem to the 'machine' codes involves, as one of its stages, the transcription of the symbols used in the formulae into digital instructions and addresses. Datum 'a' of a problem can at any time be represented by its address 'A'. Generalizing this idea, codification as a whole may be regarded as being capable even in its form of expression, of providing a representation of basic mathematical notations.

Representing operational codes by means of suggestive literal symbols,

and the addresses by means of the actual symbols of the numbers to be recorded in them, we get symbolic codification.

What the machine then has to do is to interpret the latter; and this problem is solved by the alphabetical codes being translated by a programme known as the assembly programme, based on the principle of making numerical addresses correspond to the various factors.

The various stages in the execution of a task by an electronic computer are as follows: mathematical or logical analysis of the problem, culminating in calculable formulae or patterns; preparation of a flow diagram; codification; trial run for a simple but complete numerical example; execution of the computation (final stage).

Most of the intermediate stages can be extensively mechanized. Let us consider the numerical trials, to begin with. These can be carried out automatically by the machine itself. A general-purpose programme, known as a 'tracing routing', is prepared with this aim in view, and it operates more or less as follows: When incorporated in the programme to be analysed, it allows the latter to go ahead unchanged but supplies, for each complete operation, a line of intermediate results comprising the address of the instruction being carried out, the instruction as it stands at that particular moment, the result of the operation relating to it, and the content of the accumulator and of the various corresponding registers or stores. There are some tracing programmes which supply these particulars for switch instructions only, and thus re-establish the original articulation of the programme analysed at the time of trying out.

The so-called 'diagnostic' routines help to execute work on the machine in a different way: they supply a statement on request of the computer's memories at the time of diagnosis.

Codification is probably the line of work which has proved most fruitful of all in the use of large electronic units.

The first stage, symbolic codification, has already been discussed. But modern processes are by no means limited to this 'word-for-word' transcription of a group of instructions which, after all, are only a different form of notation expressing those same instructions with which the machine is already 'familiar'.

The computer's scope is considerably enlarged once it is enabled to interpret and carry out instructions in an 'outside' language not its own.

For example, there might be prior agreement on a dictionary of new instructions which it will have to know and which will multiply its possibilities, e.g., a code for elementary functions or for floating-point or other special operations.

The general principles governing this enlargement of range are that the instructions written in the external language would be recorded in memory units as ordinary instructions, but with the user replacing the machine's

control unit by a group of interpretive programmes including look-up of a table having the instructions in the external language as the inputs, which it matches with appropriate sub-routines; and all the sub-routines concerned.

The result is what is called an interpretive routine. These codes, which may be symbolic, serve a number of purposes by making it possible: (a) for one machine to simulate another (the efficiency of the simulating machine is usually lower than that of the simulated one because of the heavier workload involved in interpreting the codes); (b) for the intrinsic possibilities of a given machine to be greatly increased. Thus a one-address fixed-point machine can easily function as if it had three addresses, floating point arithmetic and index registers. Here, too, simulation is less efficient than special pre-installed circuits would have been [29, 30].

A point to be noted, with interpretive routines, is that each instruction is carried out immediately after it has been interpreted.

At a rather more elaborate level than interpretive routines come abstractions. Instead of handling numbers represented by their addresses in the machine, they deal *en bloc* with more complex mathematical entities occupying a given space in the store, such as matrices, complex numbers and multiple-precision numbers. Our machine for simple algebraic numbers is thus transformed into an automatic tool for handling complex mathematical units.

Lastly, at the highest level of present achievement and development trends come the various types of automatic programming. They have a number of common features. They aim at supplying a directly usable programme in 'machine' language known as an 'object programme'. They are symbolic. They include instructions 'external' to the machine as well as macro-instructions, the execution of which calls for a number of sub-routines automatically inserted in the object programme.

These programming systems necessarily involve translation with the help of translator programmes known as compilers. There are two types of these at present: (a) those which are actually no more than an extension of the assembly programmes and in which a number of different macro-instructions are added to the strict instructions regarding form and content; and (b) those which, like the IBM-704 Fortran, are far less restrictive and hence far wider in scope and portray a form of algebraic language which bids fair to culminate in a common language for all machines [31, 32].

But even when this common language is found, established and put into general use, the difficulties of man-machine communication will persist unless the present punch-hole system of codifying information ceases to be a disability. The importance of research on pattern recognition in this connexion cannot be ignored.

PATTERN RECOGNITION

The ultimate object is to design and construct apparatus capable of 'reading' a written or printed page as it stands, as well as systems capable of recognizing and codifying the human voice in terms of letters and figures, thus permitting the dictation of a programme to a machine or the automatic printing of a speech.

The techniques of automatic reading being currently studied are based on a variety of principles, and the difficulties arising are considerable, not the least of them being the number of forms that manuscript or type script may take.

One method which seems to give promising results is the subdivision of the area containing the character to be identified into regular meshes [33]. These meshes are scanned and classified into three categories: black—presence of a stroke; white—absence of a stroke; grey—partial presence of a stroke or 'stain'. This distribution naturally leads to the use of a three-valued logical algebra for identification operations.

The approach adopted by other authors is much more directly linked to the study of the graphic unit itself: it has proved possible, by means of a logical calculation programme [34] on a general-purpose computer, to 'simulate' the behaviour of a reader who examines the topology of each letter. Here, as in the preceding case, the letter to be identified is presented to the machine as a matrix comprising occupied and unoccupied elements; but instead of general scanning only the nodes regarded as characterizing the form are examined before proceeding to analyse the actual outlines of the graphic unit.

Another method [35], also relying on the morphological scanning of the lines to be identified, is based on the transformation of the curve represented by the lines into voltage functions of analogous form. Identification standards are now being studied, and experiments are continuing.

Taking a rather opposite course, some specialists are attacking the problem by means of the theory of fields and potentials, and the solution of certain partial differential equations. To explain them as fairly as possible, it may be said that the character to be identified is considered as representing, within the area examined a group of lines of topographical levels [36].

General observations [37] on the recognition of written patterns were presented at the conference from the theory of information standpoint. They are of interest mainly because they help towards a better understanding of the 'prototype pattern' concept.

Processes for the identification of spoken language were discussed in only one paper [38]—which is hardly a fair reflection of the number of research workers now dealing with this problem.

Comparing linguistic analysis to chemical analysis, the authors of this

paper attempt to lay down standards and rules for 'distinguishing among speech events that are normally recognized as different by native speakers of a particular language'. In this field, also, it has proved possible to conduct experiments by means of a written programme designed for a general-purpose computer.

Special programmes have likewise made it possible to undertake some extremely interesting research work on machine learning. An already classic example is provided by the machine with a routine for playing draughts; it can never be beaten twice in the same position. This is an ordinary process of storing in the memory the best moves made by all the opponents met with.

But one of the keys to machine learning is to build machines capable of generating their own programmes [39].

Original and interesting programmes have been created by a machine designed to generate the terms of a series of numbers the first three of which are known. The machine learns by experience: the right programmes are those which succeed. It also learns to generalize simple computing procedures so as to transform them into more widely applicable methods, as in the case of the programme generating the sequence referred [40, 41].

Going beyond this elementary learning, which was itself more than anyone had dared hope for, some specialists are now translating into reality some of the most boldly visionary ideas, including the demonstrations by machine of geometrical and logical theorems [42, 43].

A great deal of emphasis is being placed on this research on heuristic methods, which alone are capable of giving a favourable bias to the search conducted by the machine [40] so as to avoid the exhaustive enumeration of all possible solutions. However, certain problems of elementary logic have been solved by the successive elimination of variables [41], thus showing the usefulness of non-heuristic methods in particular cases. In other situations [42, 43], the machine, supplied with certain general research principles and a programme for applying them, has been able to solve problems of logic of some complexity, the method of proof relying on the fact 'that a statement is provable if and only if there exists no interpretation in which the negation of the statement is valid'.

The analysis of human intellectual behaviour is the basis, of course, for work in this field, and there is obviously no question of the 'complete simulation' of reasoning processes!

The 'simulation' referred to here is the general technique whereby individual logical situations are put into a general-purpose computer by means of specific programmes. To give an idea of the question, certain problems of queuing, where the queues are of random formation, have no general theoretical solutions. All that is done, in order to solve them, is to 'simulate' the formation of queues by generating numbers at random.

The above observations need to be supplemented by the rather obvious remark that machine learning and artificial 'thinking' processes are closely linked to the concept of memory. Every improvement in devices for recording and selecting information [44, 45] has an immediate impact on the machine's capacity for learning and for pattern recognition. By establishing a formal mathematical system designed to represent the functioning of a plastic neuron 'which changes its logical properties with its activity' [46], it has been possible to prefigure circuit elements endowed to some extent with 'the large capacity and the apparent random access character of the human memory'.

Seeking in the biological field a self-adjusting study model [47], certain specialists have studied 'the general diagram of development of conditioned reflex chains'.

Thus the activity of human thought finds new outlets every day thanks to information processing machines, and the use of these automata is now being professionally extended to fields which are seemingly the least amenable to any kind of automatization. The history of research into the automatic translation of languages is a significant case in point.

AUTOMATIC TRANSLATION OF LANGUAGES

Nowhere has the idea of intuitive appreciation been so vigorously championed in opposition to the claims of technicians; and nowhere have so many mysteries been finally reduced to strict algorisms! However, much remains to be done before the first research objective—accurate translation without striving for stylistic efforts—is attained.

In his article on this subject in the present issue, E. Delavenay, Chief of Unesco's Documents and Publications Service, draws a general picture of the problems still to be solved and of the present achievements in this sphere.

The papers submitted to the conference covered: lexical problems [48, 49]; structural questions peculiar to certain languages, considered with a view to translation [50, 51, 52]; systematized programming methods [53]; machines specialized in the translation of a pair of languages [54].

The work on mechanized translation is making a substantial contribution in many cases to other fields where information processing is applied. Lexical problems are closely related to those of information selection the solution of which—according to some—depends on the success of the 'machine' translation venture.

Broadly speaking, the scientific technique of information processing which is the subject of this article is so profoundly a unity that progress in any one direction affects the entire field. In the light of the work now being done, we can only look in wonderment at the way this discipline has developed; its history began in 1944, and its 'prehistory' at the turn of the century.

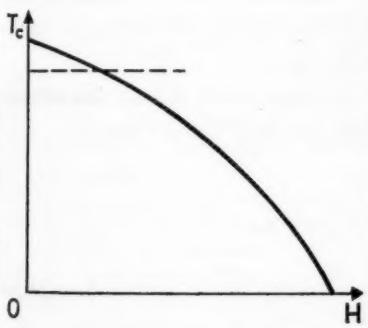
But what shall we say of the prospects as revealed by the papers outlining the techniques of the future?

FUTURE TECHNIQUES

Included under this head are all the techniques which have now reached highly interesting stages of development in the laboratories, and also the new ideas and concepts now being evolved concerning the methods of using large data processing systems. Revolutionary innovations are in preparation, and even the term 'large systems' may seem antinomic—or archaic—a few years hence.

Starting with the technological side, let us consider storage and switching circuit techniques. They depend on the most unexpected branches of physics; one of the most promising lines of advance, for example, relates to the cryogenic properties of certain substances.

Superconductivity, discovered in 1911, is a phenomenon observed in certain materials when their temperature is lowered to the neighbourhood of absolute zero: their electrical resistivity falls strictly to zero below a certain critical temperature which depends on the surrounding magnetic field. Plotting the critical temperature as the ordinates and the magnetic field as the abscissae, we obtain the curve reproduced opposite, representing the maximum values of the fields which can be applied to the substance without causing the superconductivity phenomenon to disappear. If the temperature of the substance is maintained slightly lower than the temperature for transition to a null magnetic field, its resistivity may be annulled or be given its ordinary finite value by the suppression or appearance of a relatively weak magnetic field [55].



An element of the cryogenic circuit, or 'cryotron', is constituted, in principle, by a rectilinear wire (called a 'contact') surrounded by a spiral wire (called an 'entry'), from which it is insulated. Current in the entry wire produces a magnetic field which controls the current circulating in the contact wire. Theory and calculation, confirmed by experiment, indicate that the cryotron is a circuit element from which the most outstanding performances may be expected: by combining the techniques of miniaturization [56] and refrigeration, circuits are obtained where the time unit is one thousand-millionth of a second and where the density of recorded information is in the region of 10 million binary digits per square centimetre [57].

Ultra-miniaturization techniques are all based on processes embodying the

evaporation of metals in a vacuum and condensation on an insulating substance through a microscopic grid which defines, on the deposit thus formed, a printed circuit pattern of about 1 micron.

The specialists point out that this degree of miniaturization would scarcely be justified in constructing machines as simple (!) as present-day computers. 'If a computer were reduced to the size of a packet of cigarettes, there would be only very slight advantage to reduce it still further to the size of a postage stamp.'

The magnetic properties of thin films [58] and the possibilities offered by microwave techniques and semi-conductors [59] are being exploited in work which holds out hopes of higher performance rates [60].

These technological prospects have, in fact, made us even more exacting when it comes to practical uses.

The idea of a great multitude of logical elements housed in a tiny space brings with it the tempting prospect of extrapolating the concepts of plastic neurons considered above and so providing infinite scope for the learning capacity of machines. It foreshadows a time when the titanic problems of information retrieval in a major State will perhaps be viewed with equanimity; when automatic documentation and machine translation centres will be in operation; and when hitherto insoluble problems of economic planning or weather prediction will no longer defy us.

What limits, are there, then, to all these developments? All we can answer is that it would be unsafe to forecast beyond the next twenty years!

CONFERENCE PAPERS

1. F. Ceschino and J. Kuntzmann: 'The case for reversion to the canonic form in the solution of initial condition differential problems'.
2. P. Henrici: 'Theoretical and experimental studies on the accumulation of error in the numerical solution of initial value problems for systems of ordinary differential equations'.
3. A. A. Dorodnitzin: 'The experience of the use of high-speed computers for solving partial differential equations'.
4. D. J. Evans: 'The solution of elliptic difference equations by stationary iterative processes'.
5. L. Collatz: 'Methods of computation on digital computers for partial differential equations'.
6. R. S. Varga: 'Over-relaxation applied to implicit alternating direction methods'.
7. J. H. Wilkinson: 'Rounding errors in algebraic processes'.
8. Ch. Blanc: 'Concerning the assessment of rounding errors'.
9. D. B. Gillies: 'The exact determination of the characteristic polynomial of a matrix'.
10. H. J. Maehly: 'Rational approximations for transcendental functions'.
11. G. Letellier, R. Lattès: 'Solution on a high-speed computer of a diophantine algebra problem'.
12. G. R. Parisot: 'Logarithmic programmes: their applications to the calculation of convex and, more specifically, linear programmes'.
13. C. Strachey: 'Time sharing in large, fast computers'.

14. B. J. Loopstra: 'Input and output in the x - 1 system'.
15. W. F. Schmitt and A. B. Tonik: 'Sympathetically programmed computers'.
16. J. Bosset: 'Considerations on certain logical design aspects of the Gamma-60'.
17. A. L. Leiner, W. A. Notz, J. L. Smith and R. B. Marimont: 'Concurrently operating computer systems'.
18. F. P. Brooks, Jr., G. A. Blaauw and W. Buchholz, 'Processing data in bits and pieces'.
19. I. Y. Akushsky, L. B. Emelyanov-Yaroslavsky, E. A. Klyamko, V. S. Linsky and G. D. Monakhov: 'Methods of speeding up the operation of digital computers'.
20. G. Metze and J. E. Robertson: 'Elimination of carry propagation in digital computers'.
21. H. Takahashi and E. Goto: 'Application of error-correcting codes to multi-way switching'.
22. S. Muroga: 'Logical elements on majority decision principle and complexity of their circuit'.
23. R. Vacca: 'A three-valued system of logic and its applications to base-three digital circuits'.
24. A. Svoboda: 'The numerical system of residual classes in mathematical machines'.
25. G. C. Tootill: 'The use of cyclic-permuted chain codes for digitizers'.
26. W. L. van der Poel: 'Zebra, a simple binary computer'.
27. M. Lehman: 'The specification development of a cost-limited digital computer'.
28. F. L. Bauer and K. Samelson: 'The problem of a common language, especially for scientific numerical work motives, restrictions, aims and results of the Zürich Conference on ALGOL'.
29. F. G. Duncan and E. N. Hawkins: 'Pseudo-code translation on multi-level storage machines'.
30. Y. Y. Basilevsky, Y. A. Schreider and I. Y. Akushsky: 'Methods of logical recursive and operator analysis and synthesis of automata'.
31. J. W. Backus: 'The syntax and semantics of the proposed international algebraic language of the Zürick ACM-GAMM Conference'.
32. J. Poyen and B. Vauquois: 'Suggestion for a universal language'.
33. H. Wada, S. Takahashi, T. Iijima, Y. Okumura and K. Imoto: 'An electronic reading machine'.
34. H. Sherman: 'A quasi-topological method for recognition of line patterns'.
35. W. Sprick and K. Ganzhorn: 'An analogous method for pattern recognition by following the boundary'.
36. H. Kazmierczak: 'The potential field as an aid for character recognition'.
37. Stanley Frankel: 'Information-theoretic aspects of character reading'.
38. G. W. Hughes and M. Halle: 'On the recognition of speech by machine'.
39. T. Kilburn, R. L. Grimsdale and F. H. Sumner: 'Experiments in machine learning and thinking'.
40. H. Gelernter: 'Realization of a geometry theorem-proving machine'.
41. B. Dunham, R. Fridshal and G. L. Sward: 'A non-heuristic program for proving elementary logical theorems'.
42. P. C. Gilmore: 'A program for the production of proofs for theorems derivable within the first order predicate calculus from axioms'.
43. A. Newell, J. C. Shaw and H. A. Simon: 'Report on a general problem-solving program'.
44. M. E. Stevens: 'A machine model of recall'.
45. C. N. Mooers: 'Some mathematical fundamentals of the use of symbols in information retrieval'.
46. D. G. Willis: 'Plastic neurons as memory elements'.
47. S. N. Braines, A. V. Napalkov and Y. A. Schreider: 'Analysis of the working principles of some self-adjusting systems in engineering and biology'.
48. V. E. Giuliano and A. G. Oettinger: 'Research on automatic translation at the Harvard Computation Laboratory'.
49. A. F. Parker-Rhodes and R. M. Needham: 'A reduction method for non-arithmetic data and its application to thesauric translation'.

50. K. E. Harper and D. G. Hays: 'The use of machines in the construction of a grammar and computer program for structural analysis'.
51. I. K. Belskaya: 'Machine translation methods and their application to the Anglo-Russian scheme'.
52. R. J. Solomonoff: 'A new method for discovering the grammars of phrase structure languages'.
53. V. H. Yngve: 'The COMIT system for mechanical translation'.
54. S. Takahaski, H. Wada, R. Tadenuma and S. Watanabe: 'English-Japanese machine translation'.
55. W. B. Ittner III: 'Physical characteristics of cryogenic components'.
56. D. A. Buck and K. R. Shoulders: 'An approach to microminiature printed systems'.
57. S. Methfessel, W. E. Proebster and C. Kinberg: 'Thin magnetic films'.
58. J. Raffel and D. Smith: 'A computer memory using magnetic film'.
59. J. A. Rajchman and J. W. Leas: 'Microwave solid-state techniques for high speed computers'.
60. H. E. Billing and A. O. Rvdiger: 'The possibility of speeding up computers using parametrons'.